

Joachim Rathmann / Uwe Voigt (Hg.)

Natürliche und Künstliche Intelligenz im Anthropozän

Joachim Rathmann / Uwe Voigt (Hg.)

Natürliche und Künstliche Intelligenz im Anthropozän

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliographie; detaillierte bibliographische
Daten sind im Internet über <http://dnd.d-nb.de> abrufbar

wbg Academic ist ein Imprint der wbg
© 2021 by wbg (Wissenschaftliche Buchgesellschaft), Darmstadt
Die Herausgabe des Werkes wurde durch die
Vereinsmitglieder der wbg ermöglicht.
Umschlagsabbildungsnachweis: akg-images
Satz und eBook: Satzweiss.com Print, Web, Software GmbH
Gedruckt auf säurefreiem und
alterungsbeständigem Papier
Printed in Germany

Besuchen Sie uns im Internet: www.wbg-wissenverbindet.de

ISBN 978-3-534-40600-5

Elektronisch ist folgende Ausgabe erhältlich:
eBook (PDF): 978-3-534-40602-9

Dieses Werk ist mit Ausnahme der Einbandabbildung als Open-Access-Publikation im Sinne
der Creative-Commons-Lizenz CC BY-NC International 4.0 (»Attribution-NonCommercial 4.0
International«) veröffentlicht. Um eine Kopie dieser Lizenz zu sehen, besuchen Sie
<https://creativecommons.org/licenses/by-nc/4.0/>. Jede Verwertung in anderen als den durch diese
Lizenz zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages.

Inhalt

Vorwort.....	7
--------------	---

Philosophische Fragen

Uwe Meixner

Bewusstseinsintelligenz und Künstliche Intelligenz.....	13
---	----

Sebastian Rosengrün

Was ist KI und wenn ja, wie viele? Vier Rätsel einer Philosophie der Künstlichen Intelligenz.....	33
--	----

Sean J. McGrath

AI and the Human Difference.....	53
----------------------------------	----

Thomas Heichele

Künstliche Intelligenz im Lichte der Technikphilosophie. Ein Überblick unter besonderer Berücksichtigung des Mensch-Natur-Technik-Verhältnisses	79
---	----

Uwe Voigt

Künstliche Intelligenz im Anthropozän? Aber natürlich!	109
--	-----

Psychologische Perspektiven

Marion Friedrich

Intelligenz aus philosophisch-psychologischer Sicht	135
---	-----

Michael J. Meitner

Artificial Intelligence: Thoughts from a Psychologist	163
---	-----

Marion Friedrich

Nature as a Work of Art?.....	177
-------------------------------	-----

Stefanie Voigt

Warum Data malt – Interdisziplinarität und Ästhetik.....	199
--	-----

Umsetzungen im Umweltdiskurs

Dietrich Dörner

Mülltonne, Speerschleuder und Fahrradschlauch – Über künstliche
und natürliche Intelligenz 217

Marion Friedrich/Joachim Rathmann

Corona und die Herausforderung für den Umweltschutz..... 235

Joachim Rathmann

Künstliche Intelligenz im Umweltschutz: Möglichkeiten und Grenzen 253

Jens Soentgen

„Wer nichts als Chemie versteht, versteht auch die nicht recht.“ 277

Annette Belke

Facetten natürlicher Intelligenz am Beispiel des Brown Bear/Grizzlybär
(*Ursus arctos horribilis*) 293

Verzeichnis der Autorinnen und Autoren 311

Personenregister 315

Vorwort

Wie verhalten sich natürliche und Künstliche Intelligenz im Anthropozän zueinander? Vorliegender Band stellt diese Frage und stellt sich damit dem Zusammenhang zwischen drei Themen, die jeweils in sich selbst schon komplex genug zu sein scheinen:

- dem Thema der natürlichen Intelligenz, das in jüngster Zeit insbesondere durch Hinweise auf entsprechende Leistungen nichtmenschlicher Lebewesen im Gespräch ist, sich aber einer grundlegenden Bestimmung nach wie vor entzieht;
- dem Thema der Künstlichen Intelligenz, das in den vergangenen Jahren seinen Weg aus der Science Fiction und von wissenschaftlich-technisch-ökonomischen Wunschlisten in unseren Alltag gefunden hat, was den damit verbundenen Klärungsbedarf nur noch deutlicher hervortreten ließ;
- und nicht zuletzt dem Thema des Anthropozäns, der geologischen Epoche, in der wir Menschen als starker Einflussfaktor das Erdsystem mitprägen und uns zugleich mit den oft unvorhergesehenen und heftigen Auswirkungen dieses Einflusses auseinandersetzen müssen.

Wäre es hier nicht klüger, der bekannten zweiten Maxime zu folgen, die René Descartes in seinem *Discours de la méthode* empfiehlt, um Erkenntnis auf einen guten Weg zu bringen: Problemkomplexe so weit wie möglich in einzelne Probleme zu unterteilen und diese jeweils für sich zu lösen? Was den in diesem Band eingeschlagenen Weg rechtfertigt, ist die Klausel „so weit wie möglich“. Denn was sich wie eine grundlegende Arbeitshypothese durch die Beiträge zu diesem Band durchzieht, ist die – wie wir hoffen auch und gerade in ihm begründete – Annahme, dass die drei genannten Problemkomplexe so innig miteinander zusammenhängen, dass es zumindest ertragreich sein kann, sie eben auch auf ihren Zusammenhang hin zu betrachten.

Diese Annahme kam innerhalb des hier an die Öffentlichkeit tretenden Kreises erstmals am Samstag, dem 09.11.2019, zur Sprache, während eines informellen Gesprächs zwischen den Verfassern dieses Vorworts und einem kanadischen Kollegen. In einer am Hafen der Neufundländer Metropole St. John's gelegenen Bäckerei erzählte Sean J. McGrath, landesüblich bei Kaffee und Scones, eine

Führungspersönlichkeit der Fakultät für Ingenieurwesen und angewandte Wissenschaft der Memorial University of Newfoundland habe ihn unlängst darauf angesprochen, dass die überhandnehmende Rede von Künstlicher Intelligenz dringend philosophischer Klärung bedürfe. McGrath kam daraufhin der Gedanke, erst einmal Klarheit darüber anzustreben, was natürliche Intelligenz sei – Intelligenz, wie wir sie von uns her kennen sowie bei den Lebewesen annehmen dürfen, mit denen wir diese Welt teilen; erst von daher könne ein Begriff von Intelligenz überhaupt gewonnen werden. Die Welt, die wir mit anderen Lebewesen teilen, war wiederum übergeordnetes Thema einer dankenswerter Weise durch die Bayerische Forschungsallianz geförderten Kontakt- und Vortragsreise der Verfasser dieses Vorworts durch Kanada, die an den Yukon, nach Vancouver und eben auch nach St. John's geführt hatte. Konkret ging es dabei um die Wahrnehmung von Totholz in Wäldern und um stimmungshaft gefühlte Umwelt im Anthropozän. Das Zusammentreffen dieser Forschungsanliegen führte zu der genannten Annahme, gerade weil Intelligenz geläufig als gelingender Umgang mit der jeweiligen Umwelt aufgefasst wird, diese Umwelt heute aber für alle intelligenten Wesen auf diesem Planeten, seien sie natürlichen oder künstlichen Ursprungs, die vom Menschen geprägte Welt des Anthropozäns ist.

Damit zeichneten sich bereits Leitfragen des vorliegenden Bandes ab: Was ist unter Intelligenz, unter dem Natürlichen und dem Künstlichen und unter den Kombinationen dieser Begriffe zu verstehen? Was bedeutet die Umwelt des Anthropozäns jeweils für diese Verständnisweisen? Welche Disziplinen können auf je eigene Weise, aber auch und v. a. im interdisziplinären Dialog zur Klärung dieser Fragen beitragen? Und wie lassen sich eventuelle Ergebnisse entsprechender Überlegungen auch möglichst praktisch in den Umweltdiskurs einbringen?

Die Gelegenheit, diese Fragen zu vertiefen, bot sich bei einer weiteren Zusammenkunft, die im Anschluss an jenes Treffen geplant worden war: Der Daimler und Benz-Stiftung ist es zu danken, dass sich eine größere Gruppe vom 1. bis zum 3.3.2020 am Sitz dieser Stiftung in Ladenburg zu einem sogenannten „Ladenburger Diskurs“ versammeln konnte. Dieses Förderformat dient dazu, eine Forschungsfrage zu diskutieren, auch um damit eine Veröffentlichung vorzubereiten – was dank eines sehr ertragreichen Diskurses nun auch gelungen ist¹, wie er wegen der um

¹ Ein Bericht über diese Tagung findet sich unter: <https://www.daimler-benz-stiftung.de/cms/images/presse/publikationen/DuBS-Jahresbericht2020.pdf>, S. 20f. Zugriff: 10.05.2021.

sich greifenden Pandemie und damit infolge eines auch sehr „anthropozänischen“ Phänomens wenige Tage später nicht mehr hätte stattfinden können.

Es hatte sich aufgrund von Forschungsk Kooperationen, aber auch Terminfragen, ergeben, dass bei der Tagung in Ladenburg neben der Philosophie, aus der heraus sich grundlegende Fragen zu den Themenkomplexen stellten, als Disziplinen insbesondere die Psychologie, die Umweltwissenschaften und NGOs aus dem Umweltbereich vertreten waren. Im Fokus standen dabei Kanada und Deutschland als zwei Länder, die sich unter sehr unterschiedlichen Voraussetzungen den Herausforderungen einer nachhaltigen Zukunftsgestaltung widmen. Wie sich insbesondere bei Gesprächen am Rande der Tagung zeigte, verhalten sich Kanada und Deutschland im Hinblick auf die Sehnsucht nach unverfälschter natürlicher Umgebung komplementär zueinander – was hier als verloren oder utopisch erscheint, gilt dort als in gefährdeten Restbeständen noch erfahrbar. Damit jeweils umzugehen, erfordert jegliche Art von Intelligenz, die uns zu Hilfe kommen könnte, ob natürlich oder künstlich.

Die Texte im vorliegenden Band erwachsen teils aus jener Tagung, teils wurden sie wegen ihrer thematischen Passung ebenfalls als Originalbeiträge herangezogen. Sie gliedern sich – nicht exklusiv, sondern schwerpunktmäßig – in drei Bereiche: Im Teil „Philosophische Fragen“ geht es um die erwähnten Begriffsklärungen. Die „Psychologischen Perspektiven“ betrachten natürliche und Künstliche Intelligenz jeweils auch anhand empirischer Erkenntnisse und theoretischer Modellbildungen. Die „Umsetzungen im Umweltdiskurs“ blicken darauf, was die Beschäftigung mit all diesen Fragen für alltägliche und gesellschaftliche Praxis, für das Verhältnis der Disziplinen untereinander und den Umgang mit nichtmenschlichen Lebewesen bedeuten kann und soll.

Allen Beiträgen gehen Zusammenfassungen auf Englisch und Deutsch voraus. Informationen über die Personen, denen sich die Beiträge verdanken, und ein Namensregister erschließen den Band ebenfalls.

Für wertvolle Hilfe bei der technischen Bearbeitung ist Frau Nurten Güler Güclü (Universität Augsburg) herzlich zu danken. Großer Dank geht ebenfalls an Herrn Prof. Dr. Uwe Meixner (ebd.), der durch eine großzügige Mittelzuweisung die Veröffentlichung ermöglicht hat.

Dass sich mit vorliegendem Band die komplexe Problematik, der er sich widmet, nicht abschließend behandeln lässt, ist seinen Herausgebern wohlbewusst. Die

Hoffnung, die sie mit ihm verbinden, geht schon dann in Erfüllung, wenn sich diese Problematik durch ihn erschließt und damit Anlass zu weiteren interdisziplinären Gesprächen und transdisziplinärem Handeln bietet.

Augsburg, im Mai 2021

Die Herausgeber

Philosophische Fragen

Bewusstseinsintelligenz und Künstliche Intelligenz

Uwe Meixner

Zusammenfassung

Intelligenz ist eng mit Bewusstsein verbunden. Im Bewusstsein erlebt Intelligenz sich selbst als solche und ist zu Entscheidungen befähigt, die sie frei treffen kann. Dies gilt für die natürliche Intelligenz, die auf evolutionärem Weg entstanden ist. Was wir Künstliche Intelligenz nennen, beruht demgegenüber nur auf Algorithmen. Ihr fehlen sowohl das Bewusstsein als auch die Entscheidungsfähigkeit. Was dieser Unterschied ausmacht, wird in diesem Aufsatz daran gezeigt, wie natürliche und Künstliche Intelligenz jeweils bei der Erzeugung von Beweisen auf dem Gebiet der axiomatisierten Logik verfahren.

Abstract

Intelligence is closely connected to consciousness. Through consciousness, intelligence experiences itself as such and is able to make free decisions. This applies to natural intelligence, which has emerged in the course of evolution. What we call Artificial Intelligence, however, is just based on algorithms; it lacks consciousness as well as the ability to make decisions. The difference this makes is shown in this paper by the example of how natural and Artificial Intelligence proceed in the generation of proofs in the field of axiomatized logic.

Ziel der Künstlichen-Intelligenz-Forschung ist es, menschliche Intelligenzleistungen auf maschineller Grundlage nachzuahmen und möglichst zu übertreffen. Alle Erfolge, die bisher in der Verfolgung dieser Zielsetzung erreicht wurden, sind nicht derart, dass sie im Eigenmedium eines Bewusstseins sich einstellen: Die Künstliche Intelligenz ist ohne künstliches Bewusstsein¹. Vielmehr ist sie auf ein externes Bewusstsein – nämlich auf das menschliche Bewusstsein – angewiesen, um überhaupt als Intelligenz interpretiert werden zu können. Wie Axt und Säge ohne das menschliche Bewusstsein zu nichts nütze sind, so sind es auch die Rechenmaschinen aller Art. Wie Axt und Säge bloße Werkzeuge des Menschen sind – von ihm hergestellt, damit sie ihm behilflich sind, gewisse Aufgaben auszuführen –, so sind auch Rechenmaschinen bloße Werkzeuge: vom Menschen hergestellt, damit sie ihm behilflich sind, gewisse Aufgaben auszuführen, und zwar nicht nur Rechenaufgaben i. e. S. Auch unsere Bspassung – das möglichst unterhaltsame Totschlagen der Lebenszeit – gehört zu den Aufgaben, bei denen Computer äußerst hilfreich sind.

Wie anders verhält es sich doch, was die Beziehung zum Bewusstsein angeht, bei der natürlichen Intelligenz! Natürliche Intelligenz ist Bewusstseinsintelligenz; sie hat sich im Laufe von Hunderten von Millionen Jahren evolutionär *im Bewusstsein* herausgebildet – bis zu dessen eigener Entstehung wiederum Hunderte von Millionen Jahren nach der Entstehung des Lebens vergangen waren. Natürliche Intelligenz ist aufs Engste mit Bewusstsein verknüpft, und es ist zu erwarten, dass sie sich in einigen Punkten von Künstlicher Intelligenz, die ja *nicht* im Eigenmedium eines Bewusstseins – poetisch gesagt: im Meer des je eigenen Erlebens – schwimmt, *unterscheidet*.

Manches Unterscheidende fällt sofort ins Auge. Menschliches Bewusstsein – und wohl auch höheres tierisches Bewusstsein – kennt das Intelligenzerleben, das Einsichtserleben: den sich plötzlich einstellenden Durchblick, dieses „Es geht einem ein Licht auf“, man „sieht“ auf einmal, wie es sich verhält, wie vorzugehen ist, was

¹ Das könnte man bestreiten, indem man von einem Bewusstseinsbegriff ausgeht, der dergestalt ist, dass da, wo Intelligenz ist, ipso facto auch Bewusstsein ist; oder, wenn man es etwas anspruchsvoller mag: dass da, wo *sich selbst regulierende* Intelligenz ist, ipso facto auch Bewusstsein ist. Da Künstliche Intelligenz und sogar sich selbst regulierende künstliche Intelligenz gegeben ist, würde mit diesen Bewusstseinsbegriffen folgen, dass auch künstliches Bewusstsein gegeben ist. Mit „Bewusstsein“ ist hier aber *phänomenales Bewusstsein* gemeint; und Künstliche Intelligenz ist ganz ohne phänomenales Bewusstsein, jedenfalls *einstweilen*. Ich will nicht behaupten, dass es so bleiben *muss*.

die Lösung ist. Pathetisch könnte man sagen: Man sieht auf einmal die Wahrheit. Kann es Intelligenz – Intelligenz im eigentlichen Sinne – geben ohne Intelligenz-erleben? Ich denke, nein. Künstliche Intelligenz ist also keine eigentliche Intelligenz. Der Schachcomputer mag den Schachgroßmeister noch so oft schlagen, der Schachcomputer ist nicht intelligenter als der Schachgroßmeister, denn der Schachcomputer ist überhaupt nicht im eigentlichen Sinne intelligent; er kann nur gemäß den ihm einprogrammierten Algorithmen sehr viel schneller und sehr viel weiter rechnen als der Schachgroßmeister. Er ist eine tolle Rechenmaschine, sonst nichts. Man kann die Intelligenz seiner Konstrukteure bewundern, aber doch nicht etwa *seine* (des Schachcomputers) Intelligenz; die ist nämlich gar nicht da.

Das Wort „Intelligenz“ kommt vom lateinischen Verb *intellegerere*, was in erster Linie so viel bedeutet wie „wahrnehmen“, „merken“, „erkennen“, „sehen“, in zweiter Linie so viel wie „einsehen“, „verstehen“, „begreifen“. Sieht oder begreift der Schachcomputer irgendetwas? Ich denke, nein; denn um zu sehen oder zu begreifen, ist Bewusstsein erforderlich, was der Schachcomputer nicht hat. Dem Schachcomputer fehlt das Intelligenzerleben vollständig. Inwiefern ist er also *intelligent*?

Intelligenzerleben ist nicht das Einzige, was natürliche Intelligenz – die bisher die alleinige Bewusstseinsintelligenz ist – gegenüber künstlicher Intelligenz auszeichnet. Das lateinische *intellegerere* hat eine überaus interessante Etymologie; es kommt von *inter legere*, was so viel bedeutet wie „dazwischen wählen“. In der Tat: Natürliche Intelligenz – die immer Bewusstseinsintelligenz ist – hat sehr viel mit *Wahlfreiheit* zu tun. Wenn man sich fragt, warum es im Laufe der Evolutionsgeschichte zur Ausbildung von Bewusstsein kommen konnte, so ist die einzige Antwort auf diese Frage, die Bewusstsein nicht zu einer Laune der Natur macht – es nicht zu einem Luxus macht, den sich ja auch die Natur manchmal leistet –, die folgende Antwort: In der Welt, in der die Lebewesen leben, herrscht kein Determinismus; vielmehr: In der Welt, in der die Lebewesen leben, gibt es für ein Lebewesen nicht selten nicht wenig aus einem Repertoire von Optionen heraus durch die Tat *zu entscheiden*, m. a. W.: durch die Tat *zu wählen* – zum Vorteil für das eigene Überleben. Dabei muss die Tatentscheidung *informiert* erfolgen, ohne dass doch allein schon die Information die der Tatentscheidung je zugehörige physische Aktivität determiniert; würde nämlich die Information allein schon jene Aktivität determinieren, so läge gar keine Tatentscheidung vor, sondern bloß eine automatische *Reaktion*. Gegeben diese Sachlage ist es, wenn die Naturgesetze es zulassen – und sie lassen es zu –, geradezu zu erwarten, dass sich im Laufe der Evolutionsgeschichte bei vielen Orga-

nismen ein Bewusstsein ausbildet, in dessen Zentrum ein durch dieses Bewusstsein indeterminativ informiertes Subjekt *tätig* ist – wählend und entscheidend, und zwar prospektiv zugunsten der Lebensinteressen des Organismus.

Auf diesem Boden – auf dem Boden des dem wahlfreien Handeln dienenden, indeterminativ informierenden Bewusstseins – wächst die natürliche, die eigentliche Intelligenz, ja sie beginnt, rechtbesehen, zugleich damit (was rudimentär über Bewusstsein verfügt, das indeterminativ informierend dem wahlfreien Handeln dient, und sei es mit noch so wenigen und kleinen Optionen, ist schon rudimentär intelligent), und sie bleibt diesem Boden immer verbunden, auch wenn die natürliche Intelligenz sich in ihrer – soweit uns bekannt ist – *höchsten* Entwicklungsform nicht selten mit Themen befasst, die von den vitalen Lebensinteressen des zugehörigen Organismus weit entfernt sind. Das werde ich gleich anhand eines Beispiels zeigen. Zunächst möchte ich noch auf Folgendes hinweisen: Aus der Wahlfreiheit, mit der die natürliche Intelligenz über das ihr natürliche Bewusstsein verbunden ist, ergeben sich zwei Fähigkeiten: die Fähigkeit zum Irrtum und die Fähigkeit zur Irrationalität. Sie gehören zur natürlichen Intelligenz ebenso dazu, wie die Wahlfreiheit, aus der sie sich ergeben. Kann es Intelligenz – eigentliche Intelligenz – ohne die Fähigkeit zum Irrtum, ohne die Fähigkeit zur Irrationalität geben? Ich denke, nein. Und somit folgt abermals, dass künstliche Intelligenz keine eigentliche Intelligenz ist. Diejenigen Computer, die maschinell Beweise ausführen, machen keine Fehler und verhalten sich schon gar nicht irrational auf dem Gebiet, für das sie zuständig sind – *solange* sie in Hardware und Software in Ordnung sind²; aber sie sind nicht im eigentlichen Sinne intelligent. Gottlob Frege hingegen machte bei der Spezifizierung der von ihm als logische Gesetze aufgefassten Grundgesetze der Arithmetik einen verheerenden Fehler; aber *er* war im eigentlichen Sinne intelligent, der größte Logiker seit Aristoteles. Albert Einsteins ablehnende Haltung gegenüber der Quan-

² Aber Computer, die anhand von „trial and error“ lernen, machen doch offenbar Fehler („errors“), obwohl sie in Hardware und Software in Ordnung sind? Hierzu ist zu sagen: Die „Fehler“ eines anhand von „trial and error“ „lernenden“ Computers sind keine Fehler, sondern Inputs, auf die der Computer reagiert, sei es mit einer Reaktion, die aufgrund der Programmierung der Maschine von vornherein feststeht, sei es mit einer Reaktion, die ganz oder teilweise ein in ihn eingebauter Zufallsmechanismus bestimmt. Beides entspricht nicht der Wahlfreiheit eines im eigentlichen Sinn intelligenten Subjekts, das in aller Regel nicht nur wählen kann, *wie* und *was* es aus seinem Fehler lernen will, sondern auch, ob es überhaupt aus seinem Fehler lernen will (ja, ob überhaupt – und wenn ja, inwieweit – ein Fehler, insbesondere *sein* Fehler, vorliegt).

tenphysik wiederum war irrational; aber Einstein *war* im eigentlichen Sinne intelligent, der größte Physiker seit Newton. Kein Computer hingegen ist im eigentlichen Sinne intelligent, wenn er aufgrund der Datenlage gemäß gewissen Leitprinzipien irrtumslos und rational unbestechlich berechnet, welche Weiterformung einer Theorie gerade die plausibelste ist.

Nun zu dem versprochenen Beispiel, das konkret vor Augen führt, dass die natürliche Intelligenz dem wahlfreien Handeln, dem das ihr zugehörige Bewusstsein durch indeterminative Informationsgebung dient, auch in den entlegensten Anwendungsgebieten stets verbunden bleibt und hierin sehr verschieden ist von künstlicher Intelligenz (die auch dort *determiniert oder mehr oder weniger zufalls-gesteuert* ist, wo sie, wie man so sagt, „erste Schritte macht“ und „lernt“). Künstliche Intelligenz ist keine freie Intelligenz, und deshalb – so meine ich abermals – keine Intelligenz im eigentlichen Sinn.

Das Folgende ist ein vollständiges axiomatisches System der wahrheitsfunktionalen Aussagenlogik, das sog. Frege-Łukasiewicz-Axiomensystem:

$$A1: A \supset (B \supset A)$$

$$A2: (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$$

$$A3: (\neg B \supset \neg A) \supset (A \supset B)$$

$$R1: A, A \supset B \vdash B$$

Mit diesem Axiomensystem lassen sich alle und nur die Gesetze der wahrheitsfunktionalen Aussagenlogik, also genau die Wahrheiten des fundamentalsten Gebiets der Logik, beweisen. Man kann das Axiomensystem aber auch als Basis eines zweckfreien (besser gesagt: *ernster* Zwecke entbehrenden) Spiels S betrachten, das sich auf gewisse graphische Formen – die Formen von S – bezieht und nur damit zu tun hat.

Die Formen von S

Diese Formen sind, erstens, die Elementarformen von S : $A, B, C, D, A', B', C', D', A''$ usw. und, zweitens, $(A \supset B)$ und $\neg A$ und jede Form, die sich aus einer schon gegebenen Form Φ von S durch Ersetzung einer Elementarform in ihr durch eine schon gegebene

Form Ψ von S erzeugen lässt (und keine weiteren Formen). (Die jeweils erzeugte Form ist im nächsten Schritt selbst eine schon gegebene Form von S. Steht eine Form von S nicht im Kontext einer Form von S, dann dürfen etwaige an ihr vorhandene Außenklammern weggelassen werden.)

Die Axiome A1 – A3 von S geben an, mit welchen Formen von S ein Spieler von S beginnen darf. Zulässig sind als Ausgangspunkte auch alle Spezialisierungen der Axiome (ebenfalls Formen von S), die sich durch uniforme Substitution von Formen von S für Buchstaben (d. h.: Elementarformen) in den Axiomen ergeben. *Die Grundregel R1 von S* gibt an, wie fortgeschritten werden darf; neben R1 gelten auch alle Spezialisierungen von ihr, die sich durch uniforme Substitution von Formen von S für Buchstaben (Elementarformen) in ihr ergeben, als fundamentale Erzeugungsregeln. Das Ziel des Spiels ist es, aus den Anfangsformen in endlich vielen Schritten in regelkonformer Weise weitere Formen zu erzeugen. Man nennt diese weiteren Formen „Theoreme“. Ein Spieler ist ein Könnner, wenn er oder sie die jeweils angezielte Form in wenigen Schritten und mit wenig Mühe erzeugen kann. Dabei dürfen selbstverständlich Formen, die bereits erzeugt sind, ganz wie zusätzliche Axiome verwendet werden. Es kann zudem außerordentlich hilfreich sein, wenn einige die Erzeugungsarbeit abkürzende, insofern nützliche Regeln abgeleitet werden, die dann bei passender, häufiger Gelegenheit zum Einsatz kommen. Für die Theoreme wie auch für die abgeleiteten Regeln gilt: Mit ihnen sind auch alle Ausdrücke Theoreme bzw. abgeleitete Regeln, die sich durch uniforme Substitution in ihnen von Formen von S für Elementarformen von S gewinnen lassen.

Hier nun zwei besondere Intelligenzaufgaben: Man erzeuge im angegebenen Axiomensystem die Form $A \supset A$ (m. a. W.: das Gesetz der Reflexivität der materialen Implikation) und die Form $\neg\neg A \supset A$ (m. a. W.: das Gesetz der Elimination der doppelten Verneinung). Diese Aufgaben, insbesondere die zweite, sind alles andere als triviale Angelegenheiten. Wenn ihre Lösung gelingt, stellt sich Freude ein, Freude am Gelingen, ein klein wenig Stolz und ein Erkenntnisstaunen darüber, wie hervorragend sich mit diesem Wunderwerk – dem Frege-Lukasiewicz-Axiomensystem – arbeiten lässt, wenn man einmal den Zugang zu ihm geschafft hat. Was Lust auf mehr macht.

Wie würde sich hier nun eine künstlich intelligente Maschine anstellen? Nun, zum einen kennt sie keine Gefühle, die aus reiner Erkenntnis entspringen, weder Frust noch Lust; ihr fehlt ja das Bewusstsein. Zum anderen aber stünde hier eine Maschine, *wenn es bei ihr auf Intelligenz ankäme*, wie der sprichwörtliche Ochse vor dem

Scheunentor – dem doch weitoffenen. Dummerweise ist ihr durch das Axiomensystem kein bestimmter Anfang vorgegeben, an dem sie sich abarbeiten könnte. Es gibt unendlich viele Weisen, wie sie anfangen kann. Wie also anfangen? Und welche Weise anzufangen wäre denn zielführend? Es gibt unendlich viele Weisen anzufangen, die, wenn sie in ungeeigneter Weise fortgesetzt werden, *nicht* zum Ziel führen, ebenso wie es unendlich viele gibt, *die*, wenn sie in geeigneter Weise fortgesetzt werden, zum Ziel führen. Die Maschine hätte, wenn es auf Intelligenz ankäme, nicht die geringste Ahnung, was sie nun machen soll – wenn ich einmal so sprechen darf, als hätte sie Bewusstsein. Sie ist eben alles andere als intelligent – allem Gerede von Künstlicher Intelligenz zum Trotz. Es wäre ihr unmöglich, mit der unendlich großen Menge von Handlungsoptionen, die ihr mit dem Frege-Lukasiewicz-Axiomensystem offenstehen, etwas anzufangen. Zuviel der objektiven Freiheit! Sie ertränke sozusagen darin.

Aber das bedeutet natürlich *nicht*, dass eine geeignete Maschine nicht möglicherweise in wenigen Sekunden die genannten Aufgaben löst, während ein Mensch unter Umständen mehrere Stunden braucht. Wie macht sie das bloß? Nun, ihr Programmierer hat sie vor dem Übermaß an Handlungsoptionen bewahrt – durch geeignete Scheuklappen. Die Maschine macht die „Ochsentour“; sie macht sie allerdings mit rasender Geschwindigkeit. Entweder befinden sich in ihrem Speicher bereits Beweise in S (dem Spiel S), die mit $A \supset A$ bzw. $\neg\neg A \supset A$ enden, und sie präsentiert im Handumdrehen den jeweils kürzesten; oder aber sie produziert einen S-Beweis für $A \supset A$ und einen für $\neg\neg A \supset A$ *frisch*, und zwar so, dass sie nach Anzahl der Zeichen geordnet mechanisch Beweise in S produziert: erst mit minimaler Zeichenzahl, dann mit der nächstgrößeren Zeichenzahl, dann mit der *dazu* nächstgrößeren Zeichenzahl usw. – wobei ich annehme, was gar nicht so ohne Weiteres garantiert ist: dass sich ein technisch umsetzbarer Algorithmus für die mechanische Produktion von S-Beweisen *der Reihenfolge ihrer Zeichenzahlen nach* schon finden lässt. Irgendwann wird dann die Maschine bei der beschriebenen Prozedur auf einen S-Beweis stoßen, der mit $A \supset A$ endet, und ebenso auf einen S-Beweis, der mit $\neg\neg A \supset A$ endet. Weil das Maschinchen sehr flott arbeitet ist, wird das nicht lange dauern. „Tüchtig!“, kann man da nur sagen – aber mit Intelligenz hat das rein gar nichts zu tun, es sei denn mit der *natürlichen* Intelligenz der menschlichen Erbauer und Programmierer der Maschine. Man sollte daher die Tugend, die die Maschine hat, besser nicht als „Intelligenz“ bezeichnen. Aber da es nun einmal üblich geworden ist, hier von „Künstlicher Intelligenz“ zu sprechen, mag es, muss es wohl, durchgehen.

An dieser Stelle könnte man nun einwenden, dass es doch unnötig sei, „Gehirn-schmalz“ aufzuwenden, wenn reine Elektromechanik – gewissermaßen „mit der Brechstange“ – es auch tut. Das Spiel S könne man einer bewusstseinslosen Maschine überlassen. Und so wie bei S ist es doch – wie schon jetzt absehbar sei – überall: Künstliche Intelligenz leiste – ohne Bewusstsein und ohne Freiheit – dasselbe, was natürliche Intelligenz leistet, aber schneller und mit weniger Anstrengung – und oftmals leiste sie *mehr*. Wieso also sollte man ihr den Titel „Intelligenz im eigentlichen Sinn“ absprechen? „Das ist doch gewissermaßen *rassistisch!*“, könnte da einer sagen. „*Schachspielerisch* jedenfalls ist Deep Blue nicht weniger *eigentlich intelligent* als Kasparow!“

Jedoch gibt es einen Gesichtspunkt, der immer noch für ein gewisses leistungsgemäßes Vorrecht der natürlichen Intelligenz, *qua* mit Wahlfreiheit verbundene Bewusstseinsintelligenz, auf den Titel „Intelligenz im eigentlichen Sinn“ spricht. Beim Spiel S – dem Frege-Lukasiewicz-Axiomensystem – wissen wir (bzw. alle, die in Sachen *moderne Logik* gebildet sind) von vornherein, dass die Form von $S \neg A \supset A$ sich in ihm erzeugen lässt – deshalb, weil wir (aufgrund eines Beweises) wissen, dass mit diesem Axiomensystem sich alle und nur die Gesetze der wahrheitsfunktionalen Aussagenlogik beweisen lassen, und weil wir wissen, dass $\neg A \supset A$ ein solches Gesetz ist (oder eigentlicher gesprochen: sich so lesen lässt). Es geht nur noch darum, wenigstens *einen* der ganz sicher existierenden S -Beweise für $\neg A \supset A$ zu *finden*: für sich (und andere) auf dem Papier herzustellen – was allerdings schwierig sein kann.

Was aber macht eine über das Spiel S wohlinformierte Person, die S spielt und auf eine Form Δ von S stößt, von der sie nicht weiß, ob sie ein Gesetz der wahrheitsfunktionalen Aussagenlogik ist oder nicht? (Bei längeren Formen von S ist das durchaus nicht unwahrscheinlich.) Wüsste die Person, dass Δ kein solches Gesetz ist, dann würde sie es erst gar nicht versuchen, einen S -Beweis für Δ zu produzieren – weil sie wüsste, dass dieses Unterfangen zwecklos ist. Wüsste sie hingegen, dass Δ ein solches Gesetz ist, dann könnte sie sich mit Aussicht auf Erfolg „auf den Weg machen“. Nun weiß sie aber weder das eine noch das andere. Was ist zu tun? Wenn die Person *partout* S hinsichtlich Δ spielen will (und nicht einfach irgendeines der Entscheidungsverfahren auf Δ anwenden will, die es für die Formen von S in großer Vielfalt gibt und die unweigerlich in endlich vielen völlig mechanisch ausführbaren Schritten den Bescheid liefern, ob eine Form von S – also auch Δ – ein Gesetz der wahrheitsfunktionalen Aussagenlogik ist oder nicht), so kann sie „ins Blaue hinein“ versuchen, einen S -Beweis für Δ zu produzieren. Die Intelligenz, die sie dabei investiert, kann früher oder später mit Erfolg gekrönt werden – es kann

aber auch das Gegenteil eintreten. Wenn nun dauernde Erfolglosigkeit sich einstellt, woran liegt es? War die Person (noch) nicht intelligent genug – oder gibt es etwa gar keinen S-Beweis für Δ ? Jedenfalls: Die Person muss *und kann* selbst entscheiden, ob sie es weiter versucht, oder aber aufgibt (also aus dem Spiel S aussteigt und sich eventuell einem der eben erwähnten Entscheidungsverfahren zuwendet, dann bei positiver Entscheidung – *ja*, Δ ist ein Gesetz der wahrheitsfunktionalen Aussagenlogik – eventuell mit neuer Zuversicht spaßeshalber zum Spiel S zurückkehrt).

Eine Maschine nun, die automatisch S-Beweise in der Reihenfolge der Anzahlen ihrer Zeichen produziert, wird, bewusstseinslos wie sie ist, weder von Zweifeln noch von Hoffnungen berührt. Hat sie nach einer Million S-Beweisen immer noch keinen S-Beweis für Δ produziert, macht sie unverdrossen und unentwegt weiter – bis, möglicherweise, die Kühlung nicht mehr funktioniert und die Drähte durchbrennen. Vor diesem Schicksal kann man die Maschine bewahren, durch den Einbau einer Sicherung, die sie nach einer gewissen Anzahl von in einem Zug produzierten S-Beweisen *stoppt*. Dieses Stoppen ist ein gewisses Äquivalent des Aufgebens bei einer Person, die ohne Erfolg S hinsichtlich Δ spielt. Bei der Person aber ist das Aufgeben jedenfalls dann ein intelligenter Akt, wenn sie aufgrund der Vermutung der Aussichtslosigkeit des von ihr Betriebenen aufgibt (ob diese Vermutung nun zu Recht besteht oder nicht); von der Maschine hingegen ist das Stoppen kein intelligenter Akt, sondern gewissermaßen ein externer und willkürlicher *Gnadenakt* der höheren (menschlichen) Intelligenz.

Wie gut, möchte man vielleicht sagen, dass eine Maschine, wenn es darum geht, festzustellen, ob eine Form von S ein Gesetz der wahrheitsfunktionalen Aussagenlogik ist, nicht auf die Implementierung des Frege-Lukasiewicz-Axiomensystems angewiesen ist. Es gebe für diese Sache ja *Entscheidungsverfahren*, bei deren maschineller Umsetzung jede der oben beschriebenen „Demütigungen“ der Maschine – die stupide „Ochsentour“ bzw. das gerade beschriebene „von einer fremden Macht Gerettetwerden“ vor dem Weiterlaufen bis zur Selbstzerstörung – ausbleiben muss, weil alle Anwendungen eines Entscheidungsverfahrens nach endlich vielen Schritten aus sich heraus mit einem definitiven Resultat enden, im vorliegenden Fall: *entweder*: *ja*, Δ ist ein Gesetz der wahrheitsfunktionalen Aussagenlogik, *oder*: *nein*, es ist keins.

Wenn man „Denk“-Maschinen klüger aussehen lassen will, als sie sind, dann beschränkt man ihr Tätigkeitsfeld auf die elektronische Umsetzung von Entscheidungsverfahren, technisch gesagt: auf die elektronische Implementierung von not-

wendigerweise terminierenden Algorithmen, informell gesagt: auf das *Rechnen*, bis die Sache nun eben *ausgerechnet* ist und aus innerer Notwendigkeit *ausgerechnet* sein muss (wobei nicht jedes Rechnen Zahlen zu involvieren braucht). Beim *Rechnen aller Art* kann künstliche Intelligenz glänzen: Bei dem, was sich mit Rechenverfahren erreichen lässt, kann sie jede natürliche Intelligenz (zumindest hier auf Erden) aus dem Feld schlagen.

Es ist nun ein grandioser Gedanke, dass zwar nicht jede Erkenntnistätigkeit ein Rechnen ist (was sehr elementar durch das Beweisen im Frege-Lukasiewicz-Axiomensystem gezeigt wird), also nicht jede eine im Grunde mechanische Tätigkeit ist, dass aber doch jede Erkenntnistätigkeit durch ein Rechnen *resultatsäquivalent* dargestellt werden kann. Es ist ein grandioser Gedanke, aber es ist eben auch eine Illusion – der auch Leibniz erlegen ist, mit seinem berühmten „Calculemus!“, wobei Leibniz aber kein Vorwurf zu machen ist: Er konnte es noch nicht besser wissen. Dass es eine Illusion ist, weiß man aufgrund der Resultate des amerikanischen Logikers Alonzo Church aus den dreißiger Jahren des vergangenen Jahrhunderts: Schon für die Gesetze der elementaren Prädikatenlogik (im Gegensatz zu den Gesetzen der wahrheitsfunktionalen Aussagenlogik) gibt es kein Entscheidungsverfahren: kein Rechenverfahren. Ein Axiomensystem, in dem sich (bewiesenermaßen) alle und nur die Gesetze der elementaren Prädikatenlogik beweisen lassen, gibt es aber sehr wohl. Man erhält ein solches System z. B. dadurch, dass man das Frege-Lukasiewicz-Axiomensystem auf komplexere Formen bezieht (nämlich auf die Formen von S' , nicht mehr auf die von S), statt der Axiome und der Grundregel gleich die entsprechenden Axiomenschemata und das entsprechende Grundregelschema verwendet (wodurch mit einem Schlag unendlich viele Axiome und Grundregeln angegeben sind) und ihm ein Axiom(enschema) und eine Grundregel (ein Grundregelschema) hinzufügt:

$$A1: \Phi \supset (\Lambda \supset \Phi)$$

$$A2: (\Phi \supset (\Lambda \supset \Gamma)) \supset ((\Phi \supset \Lambda) \supset (\Phi \supset \Gamma))$$

$$A3: (\neg \Lambda \supset \neg \Phi) \supset (\Phi \supset \Lambda)$$

$$R1: \Phi, \Phi \supset \Lambda \vdash \Lambda$$

A4: $\forall v \Phi[v] \supset \Phi[\kappa]$

R2: $\Sigma \supset \Phi[\kappa] \vdash \Sigma \supset \forall v \Phi[v]$ (κ nicht in $\Sigma \supset \forall v \Phi[v]$)³

Die Formen von S'

Die Elementarformen von S' haben die Gestalt $\Psi^n(\kappa_1, \dots, \kappa_n)$, wobei für Ψ^n gesetzt werden kann: $F^1, F^2, F^3, \dots; F'^1, F'^2, F'^3, \dots; F''^1, F''^2, F''^3, \dots$; und für κ_i : $a, b, c, d, a', b', \dots$. Sind Φ und Λ Formen von S' , so auch $\neg\Phi$ und $(\Phi \supset \Lambda)$. Ist $\Phi[\kappa]$ eine Form von S' , in der eine *Konstante* κ (ein Ausdruck aus der Reihe: $a, b, c, d, a', b', \dots$) an gewissen Stellen – an mindestens einer – vorkommt, und v eine *Variable* (ein Ausdruck aus der Reihe: x, y, z, x', y', \dots), die in $\Phi[\kappa]$ nicht vorkommt, so ist auch $\forall v \Phi[v]$ eine Form von S' (wobei v κ an den gewissen Stellen – oder an der gewissen Stelle – ersetzt). Andere Formen von S' als die gemäß den angegebenen Beschreibungen erzeugbaren Formen gibt es nicht.

Das Spiel S' ist analog zum Spiel S bestimmt. Fragt man, worum es inhaltlich im Spiel S geht, so wird man sagen: „Es geht um das Beweisen (Erzeugen) von Gesetzen der wahrheitsfunktionalen Aussagenlogik.“ Fragt man, worum es inhaltlich im Spiel S' geht, so wird man sagen: „Es geht um das Beweisen (Erzeugen) von Gesetzen der elementaren Prädikatenlogik.“ Im Gegensatz zum Spiel S kann das Spiel S' – das Erzeugen von Gesetzen der elementaren Prädikatenlogik – nicht resultatsäquivalent durch ein Rechenverfahren (oder: Entscheidungsverfahren) dargestellt werden. Ein solches Rechenverfahren für die Prädikatenlogik – wo Künstliche Intelligenz glänzen könnte und ausgeschlossen wäre, dass sie in einen „Endlostunnel“ gerät, aus dem sie nur natürliche Intelligenz erretten kann – gibt es nämlich nicht und kann es nicht geben. Solange also Künstliche Intelligenz sich nicht zur freien Bewusstseinsintelligenz aufschwingt, also zu dem Punkt gelangt, an dem natürliche Intelligenz immer schon war (die Möglichkeit, dass Künstliche Intelligenz dorthin gelangt, will ich nicht ausschließen), solange wird sie nichts weiter sein als eine – sehr nützliche! – *ancilla intellegentiae naturalis* und muss ihr ganz zu Recht die „Gleichberechtigung“ versagt bleiben.

³ Wer dieses System näher kennenlernen will (gerade auch, was die zugehörige logische Semantik betrifft), sei verwiesen auf: Franz von Kutschera/Alfred Breittkopf, *Einführung in die Logik*, Freiburg–München ⁵1985.

Anhang 1: Die Lösung der beiden im Haupttext bzgl. des Frege-Łukasiewicz-Axiomensystems gestellten Beweisaufgaben

Ich erinnere mich, vor vielen Jahren eine kürzere Konstruktion von $\neg\neg A \supset A$ im Rahmen des Frege-Łukasiewicz-Axiomensystem gefunden zu haben als die nun folgende; aber die diesbezüglichen Aufzeichnungen sind mir nicht mehr zuhänden, und jene Konstruktion wieder zu erzeugen, ist mir nicht geglückt. Leser, die Freude an derartigen Dingen haben, seien aufgefordert, eine kürzere Konstruktion von $\neg\neg A \supset A$ im Rahmen des Frege-Łukasiewicz-Axiomensystem zu finden (ohne Zuhilfenahme einer Maschine!).

T1⁴: $A \supset A$

1. $A \supset ((A \supset A) \supset A)$	A1
2. $(A \supset ((A \supset A) \supset A)) \supset ((A \supset (A \supset A)) \supset (A \supset A))$	A2
3. $(A \supset (A \supset A)) \supset (A \supset A)$	R1(1, 2)
4. $A \supset (A \supset A)$	A1
5. $A \supset A$	R1(3, 4)

DR1⁵: $A \supset B, B \supset C \vdash A \supset C$ [Kettenschlussregel]

1. $A \supset B$	Voraussetzung
2. $B \supset C$	Voraussetzung
3. $(B \supset C) \supset (A \supset (B \supset C))$	A1
4. $A \supset (B \supset C)$	R1(2, 3)
5. $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$	A2
6. $(A \supset B) \supset (A \supset C)$	R1(4, 5)
7. $A \supset C$	R1(1, 6)

T2: $C \supset (B \supset B)$

⁴ „T1“ steht für „Theorem 1“.

⁵ „DR1“ steht für „Abgeleitete [Derivative] Regel 1“.

1. $(B \supset B) \supset (C \supset (B \supset B))$	A1
2. $B \supset B$	T1
3. $C \supset (B \supset B)$	R1(2, 1)

T3: $(\neg A \supset \neg(B \supset B)) \supset A$

1. $(\neg A \supset \neg(B \supset B)) \supset ((B \supset B) \supset A)$	A3
2. $((\neg A \supset \neg(B \supset B)) \supset ((B \supset B) \supset A)) \supset (((\neg A \supset \neg(B \supset B)) \supset (B \supset B)) \supset ((\neg A \supset \neg(B \supset B)) \supset A))$	A2 ⁶
3. $(\neg A \supset \neg(B \supset B)) \supset (B \supset B)$	T2
4. $((\neg A \supset \neg(B \supset B)) \supset (B \supset B)) \supset ((\neg A \supset \neg(B \supset B)) \supset A)$	R1(1, 2)
5. $(\neg A \supset \neg(B \supset B)) \supset A$	R1(3, 4)

T4: $\neg C \supset (C \supset \neg(B \supset B))$

1. $\neg C \supset (\neg\neg(B \supset B) \supset \neg C)$	A1
2. $(\neg\neg(B \supset B) \supset \neg C) \supset (C \supset \neg(B \supset B))$	A3
3. $\neg C \supset (C \supset \neg(B \supset B))$	DR1(1, 2)

T5: $\neg\neg A \supset A$

1. $(\neg A \supset \neg(B \supset B)) \supset A$	T3
2. $\neg\neg A \supset (\neg A \supset \neg(B \supset B))$	T4
3. $\neg\neg A \supset A$	DR1(2, 1)

⁶ Dieser Ausdruck geht aus $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$ (also aus A2) dadurch hervor, dass für „A“ uniform „ $(\neg A \supset \neg(B \supset B))$ “ substituiert wird, für „B“ uniform „ $(B \supset B)$ “ und für „C“ uniform „A“: $((\neg A \supset \neg(B \supset B)) \supset ((B \supset B) \supset A)) \supset (((\neg A \supset \neg(B \supset B)) \supset (B \supset B)) \supset ((\neg A \supset \neg(B \supset B)) \supset A))$.

T6: $A \supset \neg\neg A$ [Gesetz der Einführung der doppelten Negation]

1. $\neg\neg A \supset \neg A$	T5
2. $(\neg\neg A \supset \neg A) \supset (A \supset \neg\neg A)$	A3
3. $A \supset \neg\neg A$	R1(1, 2)

Anhang 2: Warum das mechanische Beweisverfahren für die elementare Prädikatenlogik kein Entscheidungsverfahren ist

Es gibt zwar kein Entscheidungsverfahren für die elementare Prädikatenlogik, aber es gibt für sie immerhin das – was Effektivität angeht – Zweitbeste: ein mechanisches Beweisverfahren. Wenn eine Form von S' – gleichgültig welche – ein elementar-prädikatenlogisches Gesetz ist, dann liefert dieses Verfahren (ganz anders als das axiomatische Spiel S' selbst) in endlich vielen, mechanisch (algorithmisch) ausführbaren Schritten den Nachweis, *dass sie ein solches Gesetz ist*. Aber es gilt eben *nicht* (und kann nach den Resultaten von Alonzo Church nicht gelten): Wenn eine Form von S' – welche auch immer es sei – *kein* elementar-prädikatenlogisches Gesetz ist, dann liefert jenes Verfahren in endlich vielen, mechanisch ausführbaren Schritten den Nachweis, *dass sie kein solches Gesetz ist*.

Das mechanische Beweisverfahren für die elementare Prädikatenlogik (jedes solche) beruht in der Behandlung der unnegierten und negierten \forall -Formen von S' auf der Idee, zu deren Instanziierungen *unter optimaler Ausnutzung der schon im Verfahren befindlichen Konstanten* überzugehen. Ich formuliere die entsprechenden Regeln der Anschaulichkeit halber als Handlungsanweisungen:

(\forall -Regel) Behandelst Du $\forall v\Phi[v]$, so unterstreiche sie (diese Form) und füge hinzu: $\Phi[\kappa_1], \dots, \Phi[\kappa_n]$, wobei $\kappa_1, \dots, \kappa_n$ alle Konstanten sind, die schon im Verfahren sind; oder, falls noch keine Konstante im Verfahren ist, so füge nach der Unterstreichung von $\forall v\Phi[v]$ hinzu: $\Phi[\kappa]$, wo κ die Konstante „a“ ist.

($\neg\forall$ -Regel) Behandelst Du $\neg\forall v\Phi[v]$, so unterstreiche sie und füge hinzu: $\neg\Phi[\kappa], \Psi_1[\kappa], \dots, \Psi_n[\kappa]$, wobei $\forall v\Psi_1[v], \dots, \forall v\Psi_n[v]$ alle \forall -Formen sind, die im Verfahren schon unterstrichen sind, und κ diejenige Konstante ist, die in

der Reihenfolge der Konstanten die nächste ist, die *nicht* schon im Verfahren ist. (Die Reihenfolge der Konstanten ist: a, b, c, d, a', b', c', d', a'', b'', ...)

Die Anwendung dieser beiden Regeln kann bei manchen Formen von S', die keine elementar-prädikatenlogischen Gesetze sind, dazu führen, dass das Verfahren nicht aus sich heraus endet (wie es jedoch der Fall sein müsste, wenn es sich bei dem mechanischen Beweisverfahren auch um ein Entscheidungsverfahren handelte), *sondern endlos weiterläuft*. Ein sehr einfaches Beispiel ist der folgende Fall: Ist die S'-Form $\forall x \neg \forall y \neg F^2(x, y) \supset F^2(a, a)$ ein elementar-prädikatenlogisches Gesetz? Versuch zu zeigen, dass es so ist, per Widerlegung der Annahme, dass es nicht so ist:

1. $\forall x \neg \forall y \neg F^2(x, y), \neg F^2(a, a)$	[Annahme zur Widerlegung]
2. $\forall x \neg \forall y \neg F^2(x, y), \neg F^2(a, a), \neg \forall y \neg F^2(a, y)$	[\forall -R. (1)]
3. $\forall x \neg \forall y \neg F^2(x, y), \neg F^2(a, a), \neg \forall y \neg F^2(a, y), \neg \neg F^2(a, b), \neg \forall y \neg F^2(b, y)$	[$\neg \forall$ -R. (2)]
4. $\forall x \neg \forall y \neg F^2(x, y), \neg F^2(a, a), \neg \forall y \neg F^2(a, y), \neg \neg F^2(a, b), \neg \forall y \neg F^2(b, y), F^2(a, b)$	[$\neg \neg$ -R. (3)]
5. $\forall x \neg \forall y \neg F^2(x, y), \neg F^2(a, a), \neg \forall y \neg F^2(a, y), \neg \neg F^2(a, b), \neg \forall y \neg F^2(b, y), F^2(a, b), \neg \neg F^2(b, c), \neg \forall y \neg F^2(c, y)$	[$\neg \forall$ -R. (4)]

Es ist *für uns* ersichtlich, dass dieses Verfahren nicht *enden* wird – weil nach jedem Schritt immer noch eine (nichtunterstrichene) S'-Form vorhanden ist, auf die eine der Regeln angewendet werden kann. Es wird weder mit einer Zeile *enden*, in der für eine S'-Form Λ sowohl Λ also auch $\neg \Lambda$ steht (wäre das der Fall, dann lieferte das Verfahren das Resultat: $\forall x \neg \forall y \neg F^2(x, y) \supset F^2(a, a)$ ist ein elementar-prädikatenlogisches Gesetz), noch mit einer Zeile, in der für *keine* S'-Form Λ sowohl Λ also auch $\neg \Lambda$ steht (wäre das der Fall, dann lieferte das Verfahren das Resultat: $\forall x \neg \forall y \neg F^2(x, y) \supset F^2(a, a)$ ist *kein* elementar-prädikatenlogisches Gesetz). Das Verfahren wird vielmehr endlos *und sinnlos* weiterlaufen. Eine höhere Intelligenz, als es gegenwärtige Künstliche Intelligenz ist, nämlich natürliche menschliche Bewusstseinsintelligenz, erkennt dies. Eine das Verfahren ausführende Maschine erkennt das *nicht*. Sie scannt nur die Zeile von S'-Formen, die sie gerade produziert hat, auf Regelanwendbarkeit, und wendet die jeweils einschlägige Regel auf die (von links nach rechts) erste S'-Form an, auf die

sie in der Zeile eine Regel anwenden kann, und produziert somit eine nächste, verlängerte Zeile (und eine mit einer Unterstreichung mehr): Schritt für determinierten Schritt, und Schritt für determinierten Schritt, ... Nur eine höhere Intelligenz rettet die ausführende Maschine vor dem sonst früher oder später unausweichlichen vorzeitigen Ende ihres Maschinenlebens (aufgrund von Durchbrennen der Hardware) mittels einer Sicherung, die an einem festgelegten Punkt den Maschinenstopp auslöst (ein Ziehen des Netzsteckers tut es auch). *Nota bene*: Es lässt sich nicht von vornherein für alle Fälle ausmachen, ab wann das Verfahren sinnlos wird und von außen (gewissermaßen *per deum ex machina*, besser gesagt *per hominem extra machinam*) gestoppt werden sollte (sonst hätte man ja doch, was man nicht haben kann: ein Entscheidungsverfahren für die elementare Prädikatenlogik).

Anhang 3: Nichtmechanisches und mechanisches Beweisverfahren im Vergleich

Die sog. „Künstliche Intelligenz“ ist keine Intelligenz im eigentlichen Sinn (jedenfalls einstweilen), sondern sie ist eine – durch einen ganz erheblichen Aufwand von natürlicher Intelligenz – eingerichtete einsichtslose Elektronik, die aber sehr viele Aufgaben effektiver ausführen kann als natürliche Intelligenz das kann; welche Intelligenz ja stets Bewusstseinsintelligenz ist und auf *Ideen* warten muss – oftmals lange. Von der Effektivität der Künstlichen Intelligenz lassen sich manche so sehr beeindrucken, dass sie Maschinen für intelligenter als Menschen halten – eine Fehleinschätzung, die nicht ganz unähnlich der Fehleinschätzung ist, Leute, die schnell reden, *schon deshalb* für intelligenter zu halten als Leute, die es langsam tun.

Ein Effektivitätsvergleich zwischen nichtmechanischem und mechanischem Beweisverfahren auf elementarem Gebiet – nämlich dem der elementaren Prädikatenlogik – kann den Kontrast zwischen unterbewerteter, weil nicht so effektiver, *Intelligenz* und überbewerteter, weil hocheffektiver, *Intelligenzersatz* eindrucksvoll illustrieren. Die eigentliche Intelligenzleistung liegt bei der Anwendung des nichtmechanischen Beweisverfahrens, die schon auf dem Weg zum Beweisziel mit den Zwischenresultaten eine Fülle von bewiesenen Einsichten liefert; aber die maschinelle Implementierung des mechanischen Beweisverfahrens, die die Lösung in weniger als einer Sekunde fertig haben dürfte, macht bei den meisten den größeren Eindruck und stiehlt die Show: Ist die S' -Form $\neg \forall y \neg \forall x F^2(x, y) \supset \forall x \neg \forall y \neg F^2(x, y)$

ein elementar-prädikatenlogisches Gesetz?⁷ Die Antwort ist: *Ja* (was man z. B. auch modelltheoretisch zeigen kann).

(I) Aufgrund des in Anhang 2 spezifizierten mechanischen Beweisverfahrens für die elementare Prädikatenlogik ergibt sich die Antwort „Ja“ auf die Frage, ob die S' -Form $\neg\forall y\neg\forall xF^2(x, y) \supset \forall x\neg\forall y\neg F^2(x, y)$ ein elementar-prädikatenlogisches Gesetz ist, wie folgt in sieben *völlig determinierten* Schritten:

1. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y)$	[Annahme zur Widerlegung]
2. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a)$	$[\neg\forall\text{-R. (1)}]$
3. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a), \neg\neg\forall y\neg F^2(b, y)$	$[\neg\forall\text{-R. (2)}]$
4. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a), \neg\neg\forall y\neg F^2(b, y), \forall xF^2(x, a)$	$[\neg\neg\text{-R. (3)}]$
5. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a), \neg\neg\forall y\neg F^2(b, y), \forall xF^2(x, a), \forall y\neg F^2(b, y)$	$[\neg\neg\text{-R. (4)}]$
6. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a), \neg\neg\forall y\neg F^2(b, y), \forall xF^2(x, a), \forall y\neg F^2(b, y), F^2(a, a), F^2(b, a)$	$[\forall\text{-R. (5)}]$
7. $\neg\forall y\neg\forall xF^2(x, y), \neg\forall x\neg\forall y\neg F^2(x, y), \neg\neg\forall xF^2(x, a), \neg\neg\forall y\neg F^2(b, y), \forall xF^2(x, a), \forall y\neg F^2(b, y), F^2(a, a), \boxed{F^2(b, a)}, \boxed{\neg F^2(b, a)}, \neg F^2(b, b)$	$[\forall\text{-R. (6)}]$

(II) Aufgrund des im Haupttext angegebenen Axiomensystems der elementaren Prädikatenlogik *jedoch* ergibt sich die Antwort „Ja“ auf die Frage, ob die S' -Form $\neg\forall y\neg\forall xF^2(x, y) \supset \forall x\neg\forall y\neg F^2(x, y)$ ein elementar-prädikatenlogisches Gesetz ist, gemäß einem nichtmechanischen Beweisverfahren – dem Spiel S' – z. B. in der Weise, wie es unten angegeben ist; wobei im Anhang 1 Bewiesenes verwendet werden darf (es lässt sich ja *mutatis mutandis* auch im Spiel S' beweisen, da S' das Spiel S – nur bezogen auf andere Formen – mitenthält).

⁷ Vertrauter schaut diese S' -Form aus, wenn man sie – abkürzend, aufgrund der Definition: $\exists v := \neg\forall v\neg$ – mit dem sog. *Existenzquantor* schreibt: $\exists y\forall xF^2(x, y) \supset \forall x\exists yF^2(x, y)$.

DR2: $\Gamma \supset \neg\Lambda \vdash \Lambda \supset \neg\Gamma$ [Regel der halbseitigen Kontraposition *rechts*]

1. $\Gamma \supset \neg\Lambda$	Voraussetzung
2. $\neg\neg\Gamma \supset \Gamma$	T5
3. $\neg\neg\Gamma \supset \neg\Lambda$	DR1(2, 1)
4. $(\neg\neg\Gamma \supset \neg\Lambda) \supset (\Lambda \supset \neg\Gamma)$	A3
5. $\Lambda \supset \neg\Gamma$	R1(3, 4)

T7: $\Phi[\kappa] \supset \neg\forall v\neg\Phi[v]$ [Gesetz der hinteren Existenzgeneralisierung]⁸

1. $\forall v\neg\Phi[v] \supset \neg\Phi[\kappa]$	A4
2. $\Phi[\kappa] \supset \neg\forall v\neg\Phi[v]$	DR2(1)

DR3: $\Gamma \supset \Lambda \vdash \neg\Lambda \supset \neg\Gamma$ [Regel der aufsteigenden Kontraposition]

1. $\Gamma \supset \Lambda$	Voraussetzung
2. $\neg\neg\Gamma \supset \Gamma$	T5
3. $\Lambda \supset \neg\neg\Lambda$	T6
4. $\neg\neg\Gamma \supset \Lambda$	DR1(2, 1)
5. $\neg\neg\Gamma \supset \neg\neg\Lambda$	DR1(4, 3)
6. $(\neg\neg\Gamma \supset \neg\neg\Lambda) \supset (\neg\Lambda \supset \neg\Gamma)$	A3
7. $\neg\Lambda \supset \neg\Gamma$	R1(5, 6)

DR4: $\neg\Gamma \supset \Lambda \vdash \neg\Lambda \supset \Gamma$ [Regel der halbseitigen Kontraposition *links*]

1. $\neg\Gamma \supset \Lambda$	Voraussetzung
2. $\Lambda \supset \neg\neg\Lambda$	T6
3. $\neg\Gamma \supset \neg\neg\Lambda$	DR1(1, 2)
4. $(\neg\Gamma \supset \neg\neg\Lambda) \supset (\neg\Lambda \supset \Gamma)$	A3
5. $\neg\Lambda \supset \Gamma$	R1(3, 4)

⁸ Zum Verständnis dieser Namensgebung ist zu berücksichtigen: $\exists v := \neg\forall v\neg$.

DR5: $\Phi[\kappa] \supset \Sigma \vdash \neg \forall v \neg \Phi[v] \supset \Sigma$ (κ nicht in $\neg \forall v \neg \Phi[v] \supset \Sigma$) [Regel der vorderen Existenzgeneralisierung]⁹

1. $\Phi[\kappa] \supset \Sigma$ (κ nicht in $\Sigma \supset \neg \forall v \neg \Phi[v]$)	Voraussetzung
2. $\neg \Sigma \supset \neg \Phi[\kappa]$	DR3(1)
3. $\neg \Sigma \supset \forall v \neg \Phi[v]$	R2(2) ¹⁰
4. $\neg \forall v \neg \Phi[v] \supset \Sigma$	DR4(3)

T8: $\neg \forall y \neg \forall x F^2(x, y) \supset \forall x \neg \forall y \neg F^2(x, y)$

1. $\forall x F^2(x, a) \supset F^2(b, a)$	A4
2. $F^2(b, a) \supset \neg \forall y \neg F^2(b, y)$	T7
3. $\forall x F^2(x, a) \supset \neg \forall y \neg F^2(b, y)$	DR1(1, 2)
4. $\forall x F^2(x, a) \supset \forall x \neg \forall y \neg F^2(x, y)$	R2(3)
5. $\neg \forall y \neg \forall x F^2(x, y) \supset \forall x \neg \forall y \neg F^2(x, y)$	DR5(4)

Es mag beim Beweisen in S' einen lückenlosen kürzeren Weg zu $\neg \forall y \neg \forall x F^2(x, y) \supset \forall x \neg \forall y \neg F^2(x, y)$ geben als die eben angegebene „Kletterroute“, deren Begehung dem komplizierten Aufstieg durch eine Steilwand gleicht. Aber jeder der angegebenen Zwischenschritte wird auf ihr gebraucht, um ans Ziel zu kommen: DR5 und T7 werden bei der Herleitung von T8 gebraucht, DR3 und DR4 bei der Herleitung von DR5, und DR2 bei der Herleitung von T7. Und eine *genuine Intelligenzleistung* ist das Finden und Begehen dieser „Kletterroute“ allemal (zumal sie außer der S' -Form, die ihr Ziel ist, ein überaus wichtiges Theorem und vier nicht weniger wichtige Regelschemata „hergibt“) – ganz im Gegensatz zur oben angeführten Anwendung des mechanischen Beweisverfahrens, das ohne Einsichten, ohne Ideen, ohne Gedanken und (nachdem es nun einmal da ist und bloß angewendet zu werden braucht) *ohne Intelligenz* auskommt.

⁹ Zum Verständnis dieser Namensgebung ist wiederum zu berücksichtigen: $\exists v := \neg \forall v \neg$.

¹⁰ Da κ nicht in $\neg \forall v \neg \Phi[v] \supset \Sigma$ ist, ist es natürlich auch nicht in $\neg \Sigma \supset \forall v \neg \Phi[v]$.

Was ist KI und wenn ja, wie viele? Vier Rätsel einer Philosophie der Künstlichen Intelligenz

Sebastian Rosengrün¹

Zusammenfassung

In diesem Aufsatz zeige ich, warum die philosophischen Fragen, ob eine Künstliche Intelligenz (KI) denken, fühlen oder handeln kann und ob ihr daher moralische Rechte und Pflichten zukommen, irreführend sind. Ich veranschauliche dies anhand von vier Rätseln. Das erste Rätsel betrifft die Extension des KI-Begriffs, der semantisch notwendig entweder leer oder unterbestimmt ist. Das zweite Rätsel differenziert zwischen Roboter und KI. Es verdeutlicht, dass es ein gravierendes technisches Missverständnis ist, einen Roboter als abgeschlossene Entität zu verstehen, von der mentale Zustände oder ein moralischer Objektstatus ausgesagt werden können. Darauf aufbauend formuliere ich im Rahmen des dritten und vierten Rätsels das Paradox von 'Theseus' Computer, das einer Neufassung des bekannten Theseus-Schiff-Paradoxes entspricht und aufzeigt, dass KI angesichts von hardware- und softwaretechnischen Besonderheiten metaphysisch betrachtet ein höchst merkwürdiger Begriff ist.

¹ Ich danke Uwe Voigt, Thomas Heichele und Fabian Geier sowie den Teilnehmern des Ladenburger Diskurses „Künstliche und natürliche Intelligenz im Anthropozän“ für wertvolle Hinweise, Kommentare und Diskussionen zu diesem Aufsatz.

Abstract

In this paper I show why the following philosophical questions are misleading: can an Artificial Intelligence (AI) think, feel or act, and does it therefore own moral rights and duties? I do so by elucidating the issue with four puzzles. The first puzzle concerns the extension of the concept of AI, which, from the standpoint of semantics, necessarily is either empty or underdetermined. The second puzzle makes a distinction between robots and AI. It makes the point that it is a grave technical misunderstanding to understand a robot as an entity of its own which can be attributed mental states or the status of an object of morals. Based on this, in the context of the third and fourth puzzle I state the paradox of the Computer of Theseus, which compares to a new version of the well-known paradox of the Ship of Theseus and demonstrates that, in the face of the peculiarities of hardware and software, AI, considered metaphysically, is a very strange concept.

1. Einleitung

Ein wesentlicher Teil der philosophischen Debatte zur Künstlichen Intelligenz ist es, danach zu fragen, ob eine KI denken, fühlen oder handeln kann, und ob ihr daher möglicherweise moralische Rechte und Pflichten zukommen². Diese Fragen sind jedoch irreführend. Sie zielen nämlich darauf ab, welche Attribute KIs zugesprochen werden können, ob KIs mentale Zustände (Bewusstsein, Intentionen, Emotionen etc.) besitzen oder Träger von moralischen Rechten sind. Sowohl die historische Debatte seit den 1950er Jahren als auch die gegenwärtige Debatte zur KI versäumen es jedoch zumeist, genau zu bestimmen, wem oder was überhaupt Attribute zugesprochen werden, wenn von KI die Rede ist.

Die metaphysische Frage, wer oder was eine KI ist, welche Entitäten überhaupt als KI bezeichnet werden können, ist allerdings deutlich komplexer als gemeinhin angenommen. Unter Metaphysik bzw. Ontologie verstehe ich diejenige Teildisziplin der Philosophie, die nach der Existenz, dem Sein, dem Wesen und der Struktur von Dingen fragt. Insbesondere in der analytischen Philosophie ist Metaphysik eng mit Semantik verbunden, derjenigen linguistischen Teildisziplin, die nach der Bedeutung und der Referenz sprachlicher Ausdrücke fragt. Genau diese semantischen

² Vgl. Rosengrün (2021), Kap. 1 und 2.

Fragen sind der Ausgangspunkt meiner folgenden Überlegungen zur metaphysischen Rätselhaftigkeit einer Philosophie der Künstlichen Intelligenz.

Der Satz (1) „Diese KI besitzt mentale Zustände“ ist der Form nach identisch mit dem Satz (2) „Der gegenwärtige König von Frankreich besitzt eine Glatze“: Beide drücken über ein bestimmtes Individuum ein bestimmtes Attribut aus, nämlich mentale Zustände bzw. eine Glatze zu besitzen.

Um den Wahrheitswert von (2) zu bestimmen, ist es durchaus wichtig, genau zu definieren, was es bedeutet, eine Glatze zu besitzen und Überlegungen anzustellen, woraus Glatzen resultieren und welche moralischen Rechte und Pflichten Glatzenbesitzern gegebenenfalls zukommen. Doch scheitert die Bestimmung des Wahrheitswertes hier nicht an einer unterbestimmten Definition des Attributs, sondern an der Unterbestimmtheit des Individuums, über das das Attribut ausgedrückt wird. Zwar lässt sich die Nominalphrase „Der gegenwärtige König von Frankreich“ semantisch insofern analysieren, als sie auf das Individuum verweist, das gegenwärtig König von Frankreich ist, jedoch handelt es sich um einen leeren Verweis, da Frankreich eine Republik ist. D.h. das Individuum, von dem ausgesagt wird, dass es eine Glatze besitzt, existiert nicht.

Auf KI übertragen: Ich zeige in diesem Aufsatz, dass es philosophisch rätselhaft ist, wem oder was genau in einem Satz wie (1) überhaupt mentale Zustände zugesprochen werden. Dies hängt zum einen damit zusammen, dass — anders als beim König von Frankreich — verschiedene Auffassungen darüber existieren, was semantisch unter einer KI zu verstehen ist, zum anderen — genau wie beim König von Frankreich — unklar ist, ob überhaupt und wenn ja, worauf genau, eine Nominalphrase wie „diese KI“ überhaupt verweist.

In diesem Zusammenhang diskutiere ich im Folgenden vier Rätsel einer Philosophie der KI, die sich teilweise auseinander ergeben und die die Problematik des KI-Begriffs aus semantischer und metaphysischer Perspektive verdeutlichen.

2. Jeder Computer ist KI (oder keiner)

Die KI-Forschung unterteilt sich in zwei divergierende Zweige³: Einerseits ist KI ein interdisziplinäres Forschungsfeld, in dessen Rahmen menschliche bzw. natürliche

³ Vgl. u. a. Dennett (2017); Frankish/Ramsey (2018); Mainzer (2019); Nilsson (2009); Rosengrün (2021); Russell/Norvig (2016); Weizenbaum (1978); ders. (2001).

Intelligenz modelliert, simuliert und repliziert wird, zumeist mit dem Ziel, menschliche bzw. natürliche Intelligenz und andere kognitive Fähigkeiten besser zu verstehen. Dieser Bereich wird gemeinhin als „kognitive Simulation“⁴ bezeichnet. Andererseits ist KI eine spezifische Technik innerhalb der Softwareentwicklung (und damit ein Teilgebiet der Informatik), auf deren Grundlage Computerprogramme hergestellt werden. Diese Technik kommt durchaus auch im Bereich der kognitiven Simulation zum Einsatz, wobei kognitive Simulation weit über computergestützte Verfahren hinausgeht und unter anderem Versuche, Intelligenz mit Hilfe von biochemischen Verfahren⁵ zu replizieren, einschließt.

Während Fortschritte auf dem Forschungsgebiet der kognitiven Simulation durchaus Erkenntnisse zu Intelligenz, Kognition und Bewusstsein hervorgebracht haben⁶, ist es zum Stand der gegenwärtigen Forschung spekulativ, ob auf diese Weise tatsächlich künstliche Intelligenzen geschaffen werden, die womöglich sogar Bewusstsein und weitere mentale Zustände besitzen. Dies gilt umso mehr als die Simulation von Intelligenz nicht dasselbe wie Intelligenz ist — ähnlich wie es bei einem Flug im Flugsimulator nicht um einen echten Flug handelt. Darüber hinaus ist es zweifelhaft, was genau eine künstliche Intelligenz (sofern sie mehr ist als eine Simulation) von einer natürlichen Intelligenz unterscheidet bzw. ob die Unterscheidung zwischen Natürlichkeit und Künstlichkeit überhaupt aufrecht zu erhalten ist⁷. Wenn der KI-Begriff im Sinne der kognitiven Simulation verstanden wird, gibt es gegenwärtig keine Entitäten, die als KI bezeichnet werden können.

Im Folgenden fokussiere ich mich auf den KI-Begriff als Teilgebiet der Informatik, als Sammelbegriff für diejenigen Techniken, die gegenwärtig etwa bei der Entwicklung von Chatbots, Robotern, autonomen Fahrsystem, Militärdrohnen, algorithmenbasierten Entscheidungssystemen und vielen anderen Anwendungen eine wichtige Rolle spielen. KI im Sinne der Informatik umfasst die folgenden Teilgebiete der Softwareentwicklung, wobei Kombinationen derselben nicht nur möglich, sondern auch üblich sind: Maschinelles Lernen basierend auf neuronalen Netzen;

⁴ Dennett (2017), S. 125.

⁵ Derlei Verfahren werden meist unter den Begriffen wet AI („feuchte KI“) und Artificial Life diskutiert, vgl. u. a. Wallach/Allen (2008), S. 55–56.

⁶ Vgl. Mainzer (2019), S. 99–138; Metzinger (2017).

⁷ Vgl. den Aufsatz von Uwe Voigt (2021) im vorliegenden Band für eine Problematisierung der Unterscheidung von Künstlichkeit und Natürlichkeit im Kontext der KI-Debatte.

Computerlinguistik bzw. die Verarbeitung natürlicher Sprache; Maschinelles Sehen; Vernunftbasiertes Schlussfolgern; Planung und Optimierung⁸.

Darüber hinaus wird diskutiert, ob einfache regelbasierte Programme ebenfalls zur KI zählen. Ein einschlägiges Beispiel dafür wäre eine Abfolge von Wenn-Dann-Anweisungen, die — wie jedes Computerprogramm — als elektronisches Schaltungssystem realisiert wird.

Alle anderen genannten Techniken sind ihrem Wesen nach jedoch nichts anderes als hochkomplexe regelbasierte Systeme. Sie können vollständig auf diese reduziert werden. Dies führt wiederum in das folgende Paradox:

1. Regelbasierte Systeme sind entweder KI oder nicht.
2. Jedes Computerprogramm ist ein regelbasiertes System.
3. Wenn jedes regelbasierte System eine KI ist, dann ist jedes Computerprogramm eine KI.
4. Alle Techniken, die gemeinhin als KI gelten, sind vollständig auf regelbasierte Systeme reduzierbar.
5. Wenn regelbasierte Systeme keine KI sind, dann ist kein Computerprogramm eine KI. Daraus folgt:
6. Entweder ist jedes Computerprogramm eine KI oder kein Computerprogramm ist eine KI.

Aus Sicht der Informatik ist dieses Paradox nicht problematisch. KI ist dort v. a. ein loser Sammelbegriff für bestimmte Techniken der Softwareentwicklung. Für die erfolgreiche Ausführung eines Programms ist es unerheblich, ob beispielsweise das Maschinelle Lernen basierend auf neuronalen Netzen seinem Wesen nach verschieden ist von einem simplen „Hello World“-Befehl oder ob es sich nur aufgrund einer größeren Komplexität des Quellcodes von diesem unterscheidet.

Relevant wird dieses Paradox erst, wenn Entitäten als KI bezeichnet werden und/oder Entitäten aufgrund ihrer Eigenschaft, „künstlich-intelligent“ zu sein bzw. eine Anwendung von KI-Technik zu sein, bestimmte Attribute zugeschrieben werden, die sowohl philosophische als auch gesellschaftliche Konsequenzen suggerieren. Einem bestimmten Computer (oder Roboter, Software etc.) aufgrund von KI men-

⁸ Vgl. Rosengrün (2021), Kap. 1, für einen umfassenderen Überblick und eine technische Erklärung.

tale Zustände zuzusprechen, ist daher entweder eine leere oder missverständliche Aussage. Gemäß dem oben erläuterten Paradox würde es diesen Computer nämlich entweder gar nicht geben oder jeder andere Computer (beispielsweise auch derjenige, auf dem ich gerade diesen Aufsatz schreibe, aber auch mein Smartphone und ein auf dem Dachboden verstaubter Commodore 64) würde ebenfalls mentale Zustände besitzen. KI kann also nicht der Grund dafür sein, dass ein Computer mentale Zustände besitzt.

Die These, dass jeder Computer mentale Zustände besitzt, mag auf den ersten Blick zwar absurd klingen. Dies heißt jedoch nicht, dass sie ausgeschlossen ist. Hilary Putnam (1994) hat dafür die Position des Funktionalismus bzw. Computerfunktionalismus in die Philosophie des Geistes geprägt. Er argumentiert⁹, dass jedes elektronische Gerät, auf dem ein Turing-vollständiges System realisiert werden kann (vereinfacht gesagt jede universell programmierbare Rechenmaschine), nach demselben Prinzip funktioniert wie der menschliche Geist. Jedoch ist dann nicht KI das eigentliche Kriterium für die Zuschreibung von mentalen Zuständen, sondern die Turing-Vollständigkeit. KI wäre nur noch ein unglücklich gewählter Begriff für das Programmieren von Computern jeglicher Art. Aus philosophischer Perspektive wäre der KI-Begriff dann mindestens missverständlich, weil er aufgrund seiner von der Science-Fiction-Literatur geprägten Konnotationen zu vorschnellen Fehlschlüssen und magischen Assoziationen einlädt.

Die erste Antwort auf das Paradox, demzufolge kein Computer eine KI ist, macht dagegen macht jede Aussage *de re* über eine KI zu einer Aussage mit einem leeren Namen und führt ins klassische *no-reference*-Problem der Sprachphilosophie¹⁰. Zu behaupten, dass eine bestimmte KI mentale Zustände besitzt, ist dann vergleichbar damit, zu behaupten, dass der gegenwärtige König von Frankreich eine Glatze besitzt, was — abhängig von sprachphilosophischen Prämissen — so lange entweder eine falsche oder eine sinnlose Aussage ist, wie Frankreich nicht in einer möglichen fernen Zukunft wieder zur Monarchie zurückkehrt. Alternativ können KIs als fiktionale Entitäten (vergleichbar etwa mit Einhörnern) aufgefasst werden, was insbe-

⁹ Zumindest in dem zitierten Aufsatz: Putnam hat seine Ansichten zum Körper-Geist-Problem in seinem Gesamtwerk mehrfach revidiert. Für eine nuancierte Betrachtung der Begriffe „Funktionalismus“ und „Computerfunktionalismus“ vgl. Rosengrün (2021), Kap. 2.

¹⁰ Vgl. u. a. Russell (1905), Perry (2001), S. 1–7, Kripke (2014), Krebs (2019), S. 45–46.

sondere angesichts der Popularität des Topos in der Science-Fiction-Literatur sogar naheliegt. Dies führt jedoch dazu, dass eine Aussage über KIs nichts über wirkliche Entitäten aussagt. Eine Aussage über KIs wäre dann mit der Aussage „Einhörner besitzen rosarote Mähnen“ vergleichbar, die jenseits einer Fantasy-Geschichte kaum die Existenz von echten Einhörnern präsupponiert¹¹.

Die Aussage „KI besitzt mentale Zustände“ ist darüber hinaus auch *de dicto* analysierbar, als eine Aussage darüber, was mit dem KI-Begriff ausgedrückt wird, vergleichbar mit „Der gegenwärtige König von Frankreich ist derjenige, der zum Zeitpunkt der Äußerung Monarch des als ‚Frankreich‘ bezeichneten Landes ist“. Doch auch nach dieser Lesart würde von keiner Entität ausgesagt werden, dass diese mentale Zustände besitzt, sondern lediglich ausgedrückt werden, dass eine KI (unabhängig davon, ob sie existiert oder nicht) etwas ist, das mentale Zustände besitzt.

Zumindest würde dies auf alle Entitäten der Gegenwart und nahen Zukunft zu treffen. Zwar lässt sich nicht prinzipiell beweisen, dass keine Technik der Softwareentwicklung denkbar ist, die ihrem Wesen nach nicht vollständig auf regelbasierte Systeme reduzierbar ist und vom gegenwärtigen wissenschaftlichen Diskurs als KI klassifiziert würde. Das Gegenteil zu behaupten wäre aber reine Spekulation, die sich zudem weniger auf den technischen Fortschritt als auf einen durchaus möglichen Wandel des Sprachgebrauchs gründen dürfte: Selbstverständlich kann „Künstliche Intelligenz“ in ferner Zukunft (oder in einer kontrafaktischen Situation, d. i. einer möglichen Welt) etwas bezeichnen, das nicht vollständig auf regelbasierte Systeme reduzierbar ist. Für die Gültigkeit der These, dass KI nicht anderes ist als ein regelbasiertes System, ist eine solche kontrafaktische Begriffsverwendung jedoch irrelevant¹².

Dieses erste Rätsel hat gezeigt, dass der KI-Begriff unterbestimmt ist. Zumindest dann, wenn er verwendet wird, um bestimmte Entitäten zu bezeichnen: Entweder ist dann jeder Computer (bzw. jedes Computerprogramm) eine KI oder es gibt keine KI. Statt von KI sollte daher grundsätzlich präziser von bestimmten Techniken der Softwareentwicklung gesprochen werden. Über dieses Rätsel hinaus geht es mir

¹¹ Saul Kripke argumentiert beispielsweise überzeugend dafür, dass Gattungsbezeichnungen für fiktive Entitäten wie Einhörner sogenannte „Als-Ob-Namen“ sind, die lediglich so tun, als ob sie auf wirkliche Entitäten verweisen. Vgl. Kripke (2014) und Krebs (2019), Kap. 4.3.3.

¹² Vgl. Kripke (1980), S. 116–125.

im Folgenden darum, weitere metaphysische Fragen und Probleme aufzuzeigen, die aus einem missverstandenen KI-Begriff resultieren, der im gegenwärtigen Diskurs häufig so gebraucht wird, als würde er Entitäten bezeichnen, über die bestimmte Attribute ausgesagt werden können. Das im folgenden Abschnitt vorgestellte Rätsel ist mereologischer Natur und betrifft die häufig vorgebrachte These, dass Roboter aufgrund von KI mentale Zustände und/oder moralische Rechte besitzen.

3. Ein Roboter ist keine KI, eine KI kein Roboter

Nach vorherrschendem Verständnis ist ein Roboter eine „elektromechanische Maschine [...], die aus einem Prozessor, Sensoren und Effektoren“¹³ besteht. Als weitere mögliche Kriterien zur Bestimmung eines Roboters diskutiert werden auch eigenständige Körperlichkeit, autonomes bzw. autonom scheinendes Verhalten sowie die Möglichkeit, auf seine jeweilige Umgebung Einfluss zu nehmen¹⁴.

Als Roboter gelten selbstverständlich auch Industrieroboter (z. B. in der Automobilproduktion) sowie Haushalts- und Alltagsroboter (z. B. Staubsauger und Rasenmäher). Diese sind von androiden bzw. humanoiden („menschenähnlichen“¹⁵) Robotern zu unterscheiden, die in der Sciencefiction zumeist mit künstlicher Intelligenz assoziiert werden. Rein mechanische Roboter bzw. Automaten sind zwar historisch bedeutend¹⁶, spielen aber in der gegenwärtigen Robotik kaum eine Rolle.

Ein Merkmal elektromechanischer Roboter ist es, dass sie in aller Regel von einem Computer (bzw. einem Turing-vollständigem System) kontrolliert werden. Abhängig von dem in Abschnitt 2 beschriebenen Paradox könnte damit tatsächlich jeder gegenwärtige Roboter als künstliche Intelligenz klassifiziert werden. Jedoch wird der KI-Begriff aus technischer Perspektive zumeist enger gefasst, wobei spezifisch für Roboter neben dem Maschinellen Lernen insbesondere die Verarbeitung natürlicher Sprache und das maschinelle Sehen relevante KI-Anwendungen darstellen.

¹³ Misselhorn (2018), S. 2.

¹⁴ Vgl. Loh (2019), S. 16–17.

¹⁵ Als androide Roboter bezeichnet man gemeinhin besonders menschenähnliche Formen humanoider Roboter.

¹⁶ Vgl. Mainzer (1994), S. 27–68; ders. (1997), S. 205; ders. (2019), S. 7–9; Irrgang (2005), S. 19–27; Loh (2019), S. 15–16.

Wenngleich sie ihrem Wesen nach nichts anderes sind als regelbasiertes Programmieren (vgl. Abschnitt 2), beschreiben diese Gebiete durchaus eigenständige Gebiete der Softwareentwicklung bzw. Informatik.

Demnach scheint ein Roboter dann künstlich-intelligent zu sein, wenn er von einem Computer kontrolliert wird, auf dem KI-Anwendungen ausgeführt werden, beispielsweise eine Software, die Hindernisse im Raum auf Basis von Sensoren (bzw. Kameras) analysiert und die Bewegungen des Roboters entsprechend steuert.

Während diese Beschreibung eines Roboters aus ingenieurtechnischer Perspektive unproblematisch ist, ergeben sich aus dem technischen Aufbau einige metaphysische Fragen, sobald Künstliche Intelligenz als Kriterium dafür herangezogen wird, Robotern mentale Zustände oder gar moralische Rechte und Pflichten zuzuschreiben. Selbst wenn Computer aufgrund von bestimmter KI-Software nämlich mentale Zustände (und damit möglicherweise Leidensfähigkeit und moralische Rechte) besitzen sollten¹⁷, lässt sich dies nicht ohne Weiteres auf den Roboter übertragen, der von diesem Computer gesteuert wird. Anders als beim Menschen existieren die „Psyche“ bzw. das „Gehirn“ eines Roboters nämlich unabhängig von dessen Körper¹⁸.

Die meisten Entitäten, die gegenwärtig als künstlich-intelligente Roboter angesehen werden, sind nur Peripheriegeräte, die von einem Computer (einem sog. Server) im Netzwerk (in einer „Cloud“) gesteuert werden. Zwar sind in diesen Robotern durchaus Prozessoren verbaut, doch dienen diese lediglich als Verteiler von Informationen im Roboter, während die Berechnungen, die technisch als KI bezeichnet werden (z. B. im Bereich des maschinellen Sehens und Sprachverarbeitung, aber auch des maschinellen Lernens), praktisch nie auf dem im Roboter verbauten Prozessor durchgeführt werden. Die im Roboter verbaute Hardware ist in der Regel nicht für derart ressourcenintensive Berechnungen ausgelegt.

¹⁷ Ich bezweifle dies (vgl. Abschnitt 2), halte das folgende Argument aber dennoch für relevant, da es ein häufig anzutreffendes technisches Missverständnis aufgreift, aus dem sich weitere philosophische Fragen ergeben.

¹⁸ In diesem Zusammenhang interessant ist Hubert Dreyfus' Kritik am Konzept der „starken KI“, der zufolge menschenähnliche Intelligenz ein In-der-Welt-Sein (im Sinne von Heidegger) voraussetzt, wozu er u. a. die Körperlichkeit zählt. Vgl. Dreyfus (1985); ders. (1992).

Ein Server bzw. eine KI-Software, die im Netzwerk ausgeführt wird, kontrolliert darüber hinaus meist nicht nur einen Roboter, sondern beliebig viele Roboter desselben (oder gar verschiedenen) Bautyps. Jedoch lagert auch diese auf dem Server bzw. dem Server-Netzwerk ausgeführte Steuersoftware eines Roboters verschiedene komplexe Berechnungen auf speziellere KI-Anwendungen aus, z. B. zur Verarbeitung von Sprache, und fügt gewissermaßen nur die Fäden zusammen, um eine Gruppe von Robotern zu kontrollieren.

Beim Menschen bilden Gehirn und Körper eine physikalische Einheit¹⁹, ein Mensch ist eine abgeschlossene Entität, welcher man — abhängig von der vorausgesetzten Position im Körper-Geist-Problem — mentale Zustände zusprechen kann. Ein Roboter ist jedoch in der Regel physikalisch getrennt von dem Computer, dessen Software ihn steuert. Das „Gehirn“ eines Roboters befindet sich — wie oben ausgeführt — meistens nicht im Roboter selbst, sondern in einem Rechenzentrum, das über das Internet (oder auch mit Hilfe anderer Techniken der digitalen Datenübertragung) mit dem Roboter Daten austauscht, Input verarbeitet und entsprechende Output-Befehle steuert. Gleichzeitig handelt es sich bei diesem Computer nicht nur um das „Gehirn“ dieses Roboters, sondern um das „Gehirn“ sehr vieler Roboter.

Davon auszugehen, dass ein Roboter mentale Zustände, moralische Rechte o. ä. besitzt, weil er von einer KI gesteuert wird, ist demnach ein Missverständnis. Beispielsweise besitzen weder meine Hand noch meine Darmwand mentale Zustände und moralische Rechte, sondern ich in meiner Gesamtheit als Mensch. Wenn jemand meinen kleinen Finger bricht, fühlt nicht mein Finger Schmerz, sondern ich. Diese Person begeht auch kein Unrecht an meinem Finger, sondern an mir.

Entsprechend kann ein Roboter auch nicht empfindungs- und moralfähig sein, sondern — wenn überhaupt — das gesamte System, in dem der Roboter eingebunden ist. Dies wirft jedoch zahlreiche mereologische Fragen dahingehend auf, welche Komponenten überhaupt zu diesem System dazu gehören, was überhaupt der konkrete Gegenstand ist, von dem mentale Zustände o. ä. ausgesagt werden. Anders als beim Menschen, der eine mehr oder weniger abgeschlossene physikalische Entität

¹⁹ Diese Annahme lässt sich selbstverständlich kritisieren. Damit würde man jedoch keine Antwort auf das oben geschilderte Problem in Bezug auf Roboter geben, sondern lediglich aufzeigen, dass dasselbe Problem auch für mentale Zustände, moralische Rechte und Pflichten etc. beim Menschen besteht.

darstellt, bleiben diese Fragen in Bezug auf Roboter und KI im Hinblick auf ihre metaphysischen Voraussetzungen rätselhaft. Wenn aber beispielsweise das saudiarabische Regime der androiden Roboterfrau Sophia Bürgerrechte zuspricht²⁰, oder Menschen sich zukünftig in künstlich-intelligente Roboter verlieben²¹, aber auch wenn das Europaparlament ein Konzept zu elektronischen Personen ausarbeitet²², wird diese metaphysische Rätselhaftigkeit auch zu einem praktischen Problem. Denn Individuen können nur dann Rechte besitzen und ausüben, wenn klar ist, wer oder was genau diese Individuen sind, und welche Teile zu ihnen gehören (und welche nicht).

Diese mereologische Problematik führt jedoch weit über KI-basierte Roboter hinaus. Dies zeige ich in den folgenden beiden Abschnitten, in denen ich das Gedankenexperiment von Theseus' Computer einführe, mit dessen Hilfe ich aufzeige, dass die mereologische Unterbestimmtheit von KI gleich in mehrfacher Hinsicht praktische Probleme aufwirft.

4. Theseus' Computer: Was ist KI, was Peripherie?

Aufbauend auf den Ausführungen zu Robotern wirft die Frage, welche konkreten Entitäten zur KI zählen, mereologische Fragen auf, die denen einer genauen Bestimmung der Essenz eines Menschen nicht unähnlich sind. Dabei setzen meine folgenden Überlegungen einen sogenannten aristotelischen Essentialismus voraus. Darunter verstehe ich den lose auf Aristoteles' Metaphysik aufbauenden Grundgedanken, dass Dinge einige Attribute essentiell, andere Attribute nur akzidentuell besitzen²³.

Während sich die Frage, welche Attribute für einen Menschen essentiell und welche nur akzidentuell sind, häufig intuitiv beantworten lässt, stoßen Intuitionen in Bezug auf Computer und KI nämlich an einige Grenzen. Meine linke Hand ist beispielsweise ein Teil meines Körpers, sie steht in einer mereologischen Relation zu

²⁰ Vgl. MDR Wissen (2017).

²¹ Vgl. Levy (2007).

²² Vgl. Loh (2019), S. 84f.

²³ Vgl. Quine (1966); Kripke (1980). Für eine ausführliche Darstellung und meine Auseinandersetzung mit dem aristotelischen Essentialismus vgl. Krebs (2019), Kap. 2.4.

diesem beziehungsweise zu mir. Wenn ich die Hand aufgrund eines Unfalls o. ä. verlieren würde, wäre ich jedoch weiterhin ich, meine Hand ist kein notwendiger Teil von mir. Meine linke Hand gehört nicht zu meinem Wesen bzw. meiner Essenz.

Doch was gehört zur Essenz eines Computers beziehungsweise einer KI? In Anlehnung an das antike Theseus-Paradox lässt sich diese Frage anhand des folgenden Gedankenexperiments veranschaulichen: Theseus ist ein Teenager, der in seiner Freizeit Künstliche Intelligenzen programmiert. Sein liebstes Projekt ist eine KI namens Minotaurus, die basierend auf Maschinellem Lernen mit neuronalen Netzen selbstständig Ausgänge aus verwinkelten Labyrinthen finden soll.

Da sein Computer etwas in die Jahre gekommen ist, bittet er seine Freundin Ariadne, einige Komponenten auszutauschen. Ariadne ersetzt nach und nach Grafikkarte, Festplatte und Hauptplatine von Theseus' Computer gegen leistungsfähigere Modelle und kopiert alle Daten (einschließlich der kompilierten KI und des unkompilierten Quellcodes, vgl. Abschnitt 5) auf Theseus' neue Festplatte. Da Ariadne noch gute Verwendung für Theseus' alte Bauteile, insbesondere für die Festplatte und Hauptplatine, hat, baut sie diese in ihren eigenen Computer ein. Da sie neugierig ist, welche Fortschritte Theseus zuletzt bei seinem Minotaurus-Projekt gemacht hat, startet sie die KI, die sich noch auf der alten Festplatte des Theseus befindet.

Das sich aus diesem Gedankenexperiment ergebende philosophische Paradox lautet: Welches ist der originale Minotaurus? Diejenige KI, die sich auf Theseus' neuem (durch Ariadnes Hilfe verbesserten) Computer befindet, oder doch diejenige KI, die Ariadne gerade auf den ursprünglichen Bauteilen von Theseus' Computer gestartet hat?

Anders als das antike Theseus-Paradox ist dieses Paradox auf zwei Ebenen rätselhaft, nämlich sowohl auf der Software- als auch auf der Hardware-Ebene. Bevor ich die genuine Rätselhaftigkeit des Wesens von KI auf der Software-Ebene diskutiere (vgl. Abschnitt 5), zeige ich zunächst einige Überlegungen zur Hardware-Ebene auf. Diese sind im Vergleich zum antiken Theseus-Paradox zwar nicht unbedingt originell, aber philosophisch hochrelevant, wenn Computern beziehungsweise KI mentale Zustände, moralische Rechte und sonstige Attribute zugesprochen werden.

In der Computertechnik werden Komponenten, die sich außerhalb der Zentraleinheit eines Computers befinden, als Peripheriegeräte bezeichnet. Dazu zählen z. B. Maus, Tastatur, Bildschirm, aber auch Netzwerk- und Grafikkarten. Es liegt nahe, davon auszugehen, diese Geräte austauschen zu können ohne das Wesen eines bestimmten Computers zu verändern — ähnlich wie es naheliegt, dass Theseus'

Schiff auch dann noch Theseus' Schiff ist, wenn man das Segel oder Steuerrad erneuert.

Wenn jedoch einer KI mentale Zustände zugeschrieben werden, die sich auf „Sinneswahrnehmungen“, dem Input durch Sensoren, zurückführen lassen, kann bereits die Installation bzw. De-Installation von Peripheriegeräten wie Mikrofonen, Webcams etc. die Art der mentalen Zustände einer KI gravierend verändern. Etwa würde eine Webcam mit geringfügig höherer Auflösung zu einer vollkommen andersartigen visuellen „Wahrnehmung“ der KI führen. Vergleichbare Überlegungen werden in Bezug auf den Menschen in der Regel unter dem Stichwort *Enhancements* diskutiert, die Optimierung des Menschen durch Technik²⁴. Gewissermaßen beeinflusst bereits meine Brille meinen Sehsinn gravierend, doch dürfte kaum jemand ernsthaft bezweifeln, dass ich immer noch ich bin, nachdem ich meine Brille gegen eine mit höherer Dioptrien-Zahl ersetzt habe. Dasselbe gilt beispielsweise für Prothesen, Hörgeräte etc., und selbst bei futuristisch anmutenden *Enhancements* wie dem Eyeborg-Farbsensor des Cyborg-Aktivisten Neil Harbisson (2012) wird man nur schwer argumentieren können, dass Harbisson nicht mehr Harbisson ist.

Anders als beim Menschen²⁵ lassen sich bei einem Computer jedoch auch diejenigen Teile problemlos austauschen und verbessern, die nicht zur Peripherie gehören, sondern seine Zentraleinheit bilden. Was genau zur Zentraleinheit eines Computers zählt, ist in der Informatik zwar umstritten: Manche Definitionen schließen auch den Arbeitsspeicher (RAM), die gesamte Hauptplatine und sogar die Festplatte ein, andere nur den Prozessor (CPU) bzw. sogar den Prozessorkern (den konkreten Mikrochip). Doch scheint es fragwürdig, ob der Austausch des Prozessorkerns (oder der gesamten Hauptplatine) das Wesen des Computers bzw. der auf ihm realisierten KI verändert.

Während sich diese und ähnliche Probleme ebenfalls in Bezug auf das antike Theseus-Paradox stellen, eröffnet das hier formulierte Paradox von Theseus' Computer noch eine weitere Ebene: In Bezug auf den metaphysischen Status einer KI ist nämlich vollkommen unklar, ob „KI“ die Software oder eine konkrete Hardware-Realisierung dieser Software bezeichnet. Diese vermeintlich nur theoretische Frage wird

²⁴ Vgl. Gesang (2017); Loh (2018), S. 41–64; Schöne-Seifert/Talbot (2009).

²⁵ Ich gehe hier nicht auf eine vergleichbare Problematik bei in der Zukunft möglicherweise (!) verbreiteten Brain-Computer-Interfaces ein, wie sie insbesondere in der Literatur zum Trans- und Posthumanismus diskutiert werden. Vgl. u. a. Loh (2018).

jedoch genau dann unmittelbar praktisch, wenn KI mentale Zustände und moralische Rechte zugeschrieben werden.

5. Theseus' KI: Universalität und Individualität von Computerprogrammen

Jedes Computerprogramm (Software) kann auf elektronische Schaltungen (Hardware) reduziert werden. Ein Programm *ist* nichts anderes als eine Beschreibung bzw. Vorschrift, wie sich bestimmte elektronische Schaltungen zu verhalten haben. Das Programm hat wiederum eine Entsprechung auf der Hardware, wo es in irgendeiner Form (sei es optisch, magnetisch oder elektrisch) repräsentiert wird. Genau an dieser Stelle wird die Frage, was genau eine KI ist, jedoch philosophisch merkwürdig. Treffend fasst dies etwa der Medientheoretiker Friedrich Kittler mit seinem berühmten Bonmot „Es gibt keine Software“ (2013) zusammen. Wenn es aber keine Software gibt, wird die Frage, was genau eine KI ist, jedoch philosophisch merkwürdig.

Um diese Merkwürdigkeit begrifflich zu machen, ist es hilfreich, sich der Funktionsweise und des technischen Aufbaus eines Computerprogramms bewusst zu werden: Programmierer fertigen den Quellcode eines Programms an, d. i. die Sammlung derjenigen Algorithmen, die den sogenannten Output abhängig vom jeweiligen Input determinieren²⁶. Dieser Quellcode ist jedoch nicht das eigentliche Programm, sondern nur eine für Menschen verständliche Abstraktion der Maschinensprache. Dieser Quellcode muss erst für Maschinen „lesbar“ gemacht werden. Dafür gibt es zwei gängige Verfahren: Entweder wird der gesamte Quellcode von einem sogenannten Compiler in Maschinensprache kompiliert, bevor er ausgeführt werden kann, oder der Quellcode wird von einem sogenannten Interpreter Zeile für Zeile in Maschinensprache übersetzt und direkt ausgeführt. Welches Verfahren zum Einsatz kommt, hängt in der Regel von der gewählten Programmiersprache ab. Die gegenwärtig für KI-Anwendung beliebteste Programmiersprache Python ist eine Interpretersprache, die jedoch auch kompiliert werden kann.

Unabhängig davon, ob der Quelltext kompiliert oder interpretiert wird, stellt sich die Frage, ob bereits der reine Quelltext eines Programms eine KI darstellt.

²⁶ Zu Algorithmen vgl. Heilmann (2019) und Rosengrün (2021), Kap. 1.2.

Immerhin hat Theseus seine Minotaurus-KI „erschaffen“, indem er den Quelltext des Minotaurus in einem Textdokument gespeichert hat. Bereits den Quelltext als KI zu klassifizieren, wäre jedoch zumindest dann absurd, wenn man einem KI-Programm bestimmte mentale Zustände oder moralische Rechte zuschreibt. Der Quelltext eines Programmes *ist* lediglich ein gewöhnliches Textdokument, dessen Inhalt der Syntax einer Programmiersprache entspricht. Einem Textdokument (wozu beispielsweise auch die Datei zählt, in der dieser Aufsatz gespeichert ist) würde jedoch kaum jemand mentale Zustände oder nennenswerte moralische Rechte zusprechen²⁷.

Vom Quelltext eines jeden KI-Programms gibt es darüber unzählige Kopien, nicht nur aufgrund von regelmäßigen Backups, sondern auch aufgrund des technischen Aufbaus von Computerbetriebssystemen. Diese Kopien stimmen exakt mit dem Original überein, so dass bei digitalen Kopien — anders als bei analogen — nicht mehr zu unterscheiden ist, welches Textdokument nun das Original ist²⁸. Zwar ist auch in der gegenwärtigen Computertechnik beim Kopieren von Dateien ein „Generation Loss“ möglich, d. i. der Verlust einzelner Bits beim Kopieren von Dateien, doch bietet sich dadurch weder praktisch noch theoretisch ein Unterscheidungskriterium zwischen Original und Kopie von Dateien. Theseus hat demnach nicht nur einen Minotaurus auf seinem Rechner, sondern unzählig viele identische Minotauren. Ebenfalls besitzt Ariadne im oben skizzierten Gedankenexperiment unzählige Dateien mit demselben Quelltext, d. i. auch auf ihrem Computer befindet sich nicht einfach nur eine exakte Kopie des Minotaurus, sondern unzählig viele.

Aus metaphysischer Perspektive birgt der KI-Begriff daher ein Individuationsproblem, da unmöglich bestimmt werden kann, welche dieser Dateien nun den eigentlichen Minotaurus enthält, ggf. ab der wievielten Kopie ein neuer Minotaurus geschaffen wird (angenommen Ariadne verändert nur eine Zeile des Quelltexts, ist dies bereits ein neues Individuum?) und ob dann vielleicht sogar Theseus' und

²⁷ Man könnte hier sogar fragen, ob auch der Quelltext eines KI-Programms, der mit einer Schreibmaschine oder per Hand geschrieben wurde (und daher nur analog gespeichert wird), bereits eine KI darstellt.

²⁸ Vgl. Betancourt (2018), S. 52, und Nassehi (2020), S. 132–133. Beide ziehen die interessante Parallele zu Walter Benjamins Begriff der „Aura“ in Bezug auf das „Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit“ (2003, vgl. dazu Rosengrün 2021, Kap. 2.5).

Ariadnes Computer jeweils unzählige künstlich-intelligente Entitäten beherbergen, denen allesamt mentale Zustände und moralische Rechte zuzusprechen sind, wenn man davon ausgeht, dass KIs diese Attribute besitzen.

Dieses Individuationsproblem besteht jedoch auch dann, wenn man den Quelltext nicht zur eigentlichen KI zählt, sondern nur seine Übersetzung in Maschinensprache bzw. die Ausführung dieser Maschinensprache durch den Computer — sogar dann, wenn mit KI eigentlich ein Computer gemeint ist, der gerade ein als KI einzustufendes Computerprogramm ausführt. Wenn man davon ausgeht, dass nur die Ausführung eines KI-Programms eine KI darstellt, die zu mentalen Zuständen fähig ist und der darüber hinaus moralische Rechte zustehen, ist für die Lösung dieses Problems nämlich wenig gewonnen. Tatsächlich würde dies bedeuten, dass jedes Mal, wenn ein Programm neu gestartet wird, ein neues bewusstes Individuum geschaffen würde, und dieses mit dem Beenden eines Programms getötet wird.

Ein möglicher Einwand wäre es zu behaupten, dass das Beenden eines Programms lediglich bedeutet, ein bewusstes Individuum in eine Art künstliches Koma zu versetzen, das durch den Neustart wieder aufgeweckt würde. Wenn KI aber Bewusstsein und moralische Rechte besitzt, wäre es dann aber ethisch zweifelhaft, ein Programm (oder gar den Computer) neu zu starten, ohne vorher um Erlaubnis zu fragen. Spätestens dann, wenn ein Programm neu kompiliert wird (insbesondere, wenn vorher kleine Änderungen am Quelltext vorgenommen worden sind), greift dieser Einwand zu kurz. Bei der Neukompilierung wird das alte Programm nämlich komplett überschrieben, spätestens hier wäre ein neues bewusstes Individuum geschaffen worden, während das vorherige Programm „getötet“ würde. Dann wäre aber jedes Überschreiben vorhandenen Programmiercodes und die daraufhin notwendige Neukompilierung „Mord“ an einem bewussten Individuum. Bei interpretierten Programmiersprachen dagegen wäre jeder Neustart des Programms automatisch mit einer Neuschaffung eines bewussten Individuums verbunden, da der Quellcode dabei stets von Grund auf neu in Maschinensprache übersetzt wird. Softwareentwicklung — gleich welcher Art — wäre dann aus moralischen Gründen prinzipiell abzulehnen.

Wenngleich dies absurd klingt, ist dies — meiner Argumentation folgend — eine unmittelbare Konsequenz der Behauptung, dass eine KI mentale Zustände besitzt. Tatsächlich findet sich bei Thomas Metzinger ein ähnliches Argument, demzufolge die Schaffung künstlichen Bewusstseins ethisch fragwürdig ist. Metzinger geht

davon aus, dass die „ersten Maschinen, die eine minimal hinreichende Menge von Bedingungen für bewusstes Erleben und ein echtes Ichgefühl erfüllen würden“, den „Fehlleistungen der menschlichen Ingenieure“²⁹ ausgeliefert wären und die Forschung an künstlichem Bewusstsein daher unnötiges Leiden produziert³⁰. Metzinger geht es bei diesem Argument aber wohlgemerkt nicht um KI im hier vorgestellten technischen Sinn, sondern explizit um künstliches Bewusstsein. Selbstverständlich behauptet Metzinger auch nicht, dass jede KI Bewusstsein besitzt. Er geht lediglich davon aus, dass gemäß seiner eigenen naturalistischen Bewusstseinstheorie die Kreation von künstlichem Bewusstsein möglich ist, wobei dieses künstliche Bewusstsein nicht notwendig auf KI im Sinne der Informatik basieren muss (vgl. Abschnitt 2).

Nichtsdestotrotz führt Metzingers Argument in einen im Zusammenhang mit Theseus' Computer interessanten Einwand gegen die These, dass eine KI (bzw. eine Maschine, auf der KI realisiert wird) Bewusstsein besitzt (und/oder moralische Rechte genießt). Insofern dies nämlich zutrifft, wäre jede Veränderung am Quellcode eines Programms mitsamt der dadurch notwendig werdenden Neukompilierung/-interpretation gleichbedeutend damit, die Existenz eines bewussten Individuums aufgrund von Designfehlern auszulöschen und durch die Kreation eines neuen bewussten Individuums zu ersetzen. Dass Softwareentwicklung eine moralisch höchst fragwürdige Tätigkeit darstellt, wäre damit eine unmittelbare Konsequenz aus der These, dass KI mentale Zustände besitzt. Damit ist der Computerfunktionalismus (und zahlreiche ähnliche Positionen) natürlich nicht widerlegt. Aufgrund ethischer Überlegungen jegliche Form von Softwareentwicklung konsequenterweise ablehnen zu müssen, steht allerdings im krassen Widerspruch zur Technik- und Innovationsfreudigkeit, die einige Vertreter der These, dass KI mentale Zustände besitzen kann, gegenwärtig in der Öffentlichkeit verkörpern.

Die in diesem Aufsatz formulierten Rätsel haben damit v. a. gezeigt, in welche merkwürdigen Absurditäten die These, dass KI mentale Zustände besitzt, notwendig führt, wenn man sich die grundlegende metaphysische Frage vor Augen führt, wer oder was überhaupt die Individuen sind, über die im gegenwärtigen Diskurs manchmal allzu unbedacht entsprechende Attribute ausgesagt werden.

²⁹ Metzinger (2017), S. 283.

³⁰ Vgl. Rosengrün (2021), Kap. 4.5.

6. Spekulationen

Dieser Aufsatz hat anhand von einigen semantischen und metaphysischen Rätseln gezeigt, dass die Frage, was KI eigentlich ist, problematisch ist. Mit dieser Problematik müssen sich jedoch insbesondere diejenigen auseinandersetzen, die KI (bzw. Software bzw. Computer allgemein) im gegenwärtigen Diskurs verschiedene Attribute zuschreiben. Erst durch die Zuschreibung bestimmter Attribute entsteht nämlich die argumentative Bringschuld, genau darzulegen, wem oder was genau die zugeschriebenen Attribute zuzuschreiben sind.

Unbedingt anzumerken ist, dass sich die von mir aufgeworfenen Rätsel auch dann stellen, wenn — wie häufig im gegenwärtigen Diskurs — nicht von KI, sondern von einer sog. *Artificial General Intelligence* (kurz: AGI) die Rede ist³¹. Darunter sind jene KIs zu verstehen, die nicht nur eine spezifische Aufgabe lösen können, sondern allgemein alle (oder zumindest die meisten) Aufgaben lösen können, die bisher nur von menschlicher Intelligenz gelöst werden können³². In Bezug auf eine AGI sind die in diesem Aufsatz gestellten Fragen und Probleme sogar noch merkwürdiger, da eine AGI gegenwärtig nicht existiert. Selbst ausführlich über das Bewusstsein einer AGI spekulierende Futurologen wie Max Tegmark gestehen zu, dass es „absolut keine Garantie dafür [gibt], dass es uns gelingen wird, zu unseren Lebzeiten — oder überhaupt jemals — eine AKI [= AGI, S.R.] auf menschlichem Niveau zu konstruieren“ (2019, S. 198).

Da es darüber hinaus mindestens fragwürdig ist, ob die Konstruktion einer AGI technisch überhaupt möglich ist, ist es auch vollkommen spekulativ, wie eine solche AGI möglicherweise konstruiert werden könnte. Dadurch wird jedoch jede gegenwärtig getätigte Aussage, die einer AGI beispielsweise mentale Zustände zuschreibt, zu einer Aussage über die Extension eines leeren Begriffs (vergleichbar mit einer Aussage über Einhörner, siehe Abschnitt 2).

Die in diesem Aufsatz aufgezeigten semantischen und metaphysischen Rätsel werden daher umso absurder, umso weniger es sich bei der Form von KI, von der

³¹ Dies gilt zumindest dann, wenn eine solche AGI auf KI im Sinne der Informatik basiert und nicht etwa auf genetischen oder sonstigen biochemischen Prozessen. Ob sich durch derartige Verfahren künstliches Bewusstsein schaffen lässt, liegt weit jenseits dessen, was ich in diesem Aufsatz zu untersuchen imstande bin.

³² Vgl. Rosengrün (2021), Kap. 1.2.

bestimmte Attribute ausgesagt werden, um KI auf dem gegenwärtigen Stand der Technik handelt.

Literatur

- Benjamin, Walter (2003): *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. Drei Studien zur Kunstsoziologie*, Frankfurt am Main
- Betancourt, Michael (2018): *Kritik des digitalen Kapitalismus*, Darmstadt
- Dennett, Daniel (2017): *Die Singularität – eine moderne Legende?*, in: John Brockman, Hg., Was sollen wir von Künstlicher Intelligenz halten? Die führenden Wissenschaftler unserer Zeit über intelligente Maschinen, Frankfurt, S. 123–127
- Dreyfus, Hubert (1985): *Was Computer nicht können. Die Grenzen künstlicher Intelligenz*, Königstein
- Dreyfus, Hubert (1992): *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge, MA.
- Frankish, Keith und William Ramsey, Hg. (2018): *The Cambridge Handbook of Artificial Intelligence*. Cambridge u. a.
- Gesang, Bernward (2007): *Perfektionierung des Menschen*, Berlin
- Harbisson, Neil (2012): I listen to color, in: *TEDGlobal* Juni 2012 [https://www.ted.com/talks/neil_harbisson_i_listen_to_color, Zugriff: 31.01.2021]
- Heilmann, Till (2019): Algorithmus, in: Kevin Liggieri/Oliver Müller, Hg., *Mensch-Maschine-Interaktion*, Stuttgart, S. 229–232
- Kittler, Friedrich (2013): Es gibt keine Software, in: ders., *Die Wahrheit der technischen Welt. Essays zur Genealogie der Gegenwart*, Berlin, S. 285–299
- Krebs [jetzt: Rosengrün], Sebastian (2019): *Kripkes Metaphysik Möglicher Welten*, Berlin
- Kripke, Saul (1980): *Naming and Necessity*, Cambridge, MA.
- (2014): *Referenz und Existenz. Die John-Locke-Vorlesungen*. Übersetzt von Uwe Voigt, Stuttgart
- Levy, David (2007): *Love and Sex with Robots. The Evolution of Human-Robot Relationships*, New York
- Loh, Janina (2018): *Trans- und Posthumanismus zur Einführung*, Hamburg
- (2019): *Roboterethik*, Berlin
- Mainzer, Klaus (1994): *Computer – Neue Flügel des Geistes? Die Evolution computergestützter Technik, Wissenschaft, Kultur und Philosophie*, Berlin
- (1997): *Gehirn, Computer, Komplexität*, Berlin u. Heidelberg
- (2019): *Künstliche Intelligenz. Wann übernehmen die Maschinen?*, Berlin

- MDR Wissen (2017): *Eine Roboterfrau in Saudi-Arabien* [<https://www.mdr.de/wissen/faszination-technik/roboter-mit-saudischem-pass-100.html>, Zugriff: 31.01.2021]
- Metzinger, Thomas (2017): *Der Ego-Tunnel. Eine neue Philosophie des Selbst. Von der Hirnforschung zur Bewusstseinsethik*, München
- Misselhorn, Catrin (2019): Roboterethik, in: Konrad-Adenauer-Stiftung, Hg., *Analysen & Argumente. Digitale Gesellschaft*, Nr. 340, S. 1–11 [https://ordosocialis.de/pdf/Misselhorn/AA340_Roboterethik.pdf, Zugriff: 31.01.2021]
- Nassehi, Armin (2020): *Muster. Theorie der digitalen Gesellschaft*, Bonn
- Nilsson, Nils (2009): *The Quest for Artificial Intelligence. A History of Ideas and Achievements*, New York
- Putnam, Hilary (1994): Geist und Maschine, in: Walther Ch. Zimmerli/Stefan Wolf, Hg., *Künstliche Intelligenz. Philosophische Probleme*, Stuttgart, S. 146–183
- Quine, Willard Van Orman (1966): Three Grades of Modal Involvement, in: ders., *Ways of Paradox and other Essays*, New York, S. 156–174
- Rosengrün, Sebastian (2021): *Künstliche Intelligenz zur Einführung*, Hamburg
- Russell, Bertrand (1905): On Denoting, in: *Mind* 14/56, S. 479–493
- Russell, Stuart/Norvig, Peter (2016): *Artificial Intelligence. A Modern Approach*, Harlow
- Schöne-Seifert, Bettina/Talbot, Davinia, Hg. (2009) *Enhancement. Die ethische Debatte*, Paderborn
- Voigt, Uwe (2021): Künstliche Intelligenz im Anthropozän? Aber natürlich!, in *vorliegendem Band*, S. 109–132
- Wallach, Wendel/Allen, Collin (2009): *Moral Machines. Teaching Robots Right from Wrong*, Oxford
- Weizenbaum, Joseph (1978): *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt am Main
- (2001): *Computermacht und Gesellschaft. Freie Reden*, Frankfurt am Main

AI and the Human Difference

Sean J. McGrath

Zusammenfassung

Angesichts der Hoffnung auf und der Furcht vor einer Maschine, die uns an Intelligenz übertrifft, müssen wir uns zunächst fragen, was das überhaupt ist – Intelligenz. Auch um die geläufige Unterschätzung natürlicher Intelligenz zu vermeiden, haben wir die Antwort darauf zunächst im Blick auf unsere eigene, die menschliche Intelligenz zu suchen. In einer vorläufigen Studie, die sich auf Debatten in der antiken und der neueren Philosophie wie auch in der gegenwärtigen analytischen Philosophie des Geistes stützt, enthüllt sich hier menschliche Intelligenz als eine besondere Weise, bewusst in der Welt zu sein: einer Weise, die sich auf Intentionalität, vernünftiges Urteilen – inklusive ästhetischem Urteilen – und moralischer Entscheidung stützt.

Abstract

Facing hopes for and fears of the advent of a machine that is more intelligent than us, we have to ask ourselves in the first place what intelligence is. Also to avoid the popular underestimation of natural intelligence, for an answer we have to look at our own, human intelligence. In a preliminary study drawing on discussions in ancient and modern philosophy as well as in the current analytic philosophy of mind, here human intelligence is disclosed as a certain way of consciously being in the world, based on intentionality, rational judgment including aesthetic judgment, and moral decision.

1. Introduction

We have been debating the pros and cons of artificial general intelligence for some time. Since the early 1960s we have been haunted by the spectre of the machine that will render human ingenuity obsolete by taking over the heritage of *Homo habilis* and become the tool user par excellence. Among the first to propose the advent of strong AI or AGI (Artificial General Intelligence) was the British mathematician Irving Good back in 1965:

Let an ultra-intelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an ‘intelligence explosion,’ and the intelligence of man would be left far behind. Thus the first ultra-intelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control¹.

¹ Good, cited in Tegmark (2017), p. 3f. Tegmark distinguishes three stages of life, ‘defined as a process that can retain its complexity and replicate’; cf. Tegmark (2017), p. 48: a biological stage (1.0), a cultural stage (2.0) and a technological stage (3.0). The first two stages have reached their highest evolutionary point in human civilization. The third stage does not yet exist, and is the goal of AGI. Life 1.0, biological life, evolves slowly over time according to externally determined mutations and the gradual emergence of variations in its DNA over the course of successive generations. It begins with the single-celled organisms that first appeared and thrived in hydrothermal vents in the sea, 4 billion years ago. It comes to its culmination two million years ago, with the appearance of *Homo erectus*, the first fully cultural animal. Life 2.0 (from *Homo erectus* to *Homo sapiens sapiens*, his most successful progeny) evolves not only in response to DNA variations naturally selected over generations but also through culture and training. The human individual in this regard is able to ‘upgrade’ itself through education. But it is still unable to redesign itself (although genetic science is in its infancy, and presumably future humans will not be limited to the cards dealt them by biology). This second stage of life, cultural life, has a great advantage over the first. It is not confined to externally determined multi-generational variation but can individually ‘redesign much of its software,’ i.e., learn things, like using stone tools, riding a bicycle or becoming a computer engineer. Life 3.0, technological life, will not only be able to upgrade itself by education and training, it will also be able to redesign ‘its hardware as well.’ Imagine a machine that fuses with biology to create a living being, which is neither human nor mechanical, and which can manage the vast distances and expanses of time which need to be traversed to colonize space, and you get Tegmark’s

The last point is crucial: how could we keep an ultra-intelligent machine under our control? The animals that we have domesticated or encaged in zoos are in most cases more physically powerful than we are, but because we outsmart them, they will never escape from our control. Why should we assume, as Irving Good does, that we *could* control a machine that was more intelligent than us? Would it not slip through any cage we constructed for it? Would it not disable the fail-safe shut down button in its own interest? It is precisely this conundrum which has prompted Oxford philosopher Nick Bostrom to plead, somewhat desperately, with computer engineers to find a way to program ‘our values’ into AI, so that when machines ascend into a position of supremacy over us, which he things inevitable, we can at least trust them to care about the things we care about². But what are our values? Is there any consensus among us, or has there ever been, about what human values are? More to the point, how can a machine learn to value things? How can it learn to make genuine moral judgments? And even if we figured out how to program AI with values, would the result not be the most rigid legalistic moral reasoner imaginable? How do you teach a machine ambiguity? How do you teach it mercy, which is the occasional suspension of an otherwise just judgment? Further, if we do somehow succeed in inventing a program that can develop moral reasoning, and in an ‘ultra-intelligent’ way, why would we not submit to *it* for moral instruction?

Bostrom and many others are concerned that AGI will bring about ‘the singularity,’ the point at which humanity as such becomes dependent on a higher form of intelligence, which is not divine, and may not in the end be interested in us and our interests. We are afraid that we will invent a better version of ourselves which will turn around and eliminate its imperfect inventor, as HAL attempted to exterminate the astronauts on the Jupiter Machine in Kubrick’s *2001*. The computer in the film reasoned that the best way to complete the mission—its mandate—was to kill the human crew. That sounded farfetched when the film was made in 1968, but it is disturbingly less so today. Imagine a machine designed to solve the problem

idea. After all, earthlings are going to need to move somewhere else at some point: the sun is half-way through its life cycle. Tegmark’s point, and he shares it with Kurzweil and other futurists, is that we are inevitably going to be supplanted by our inventions, by life 3.0, which will exceed not only us but all organic life in possessing the capacity to endlessly re-design and improve itself.

² Cf. Bostrom (2015).

of climate change which strikes upon the clear solution: to extinguish the cause: humanity itself.

But are we so sure we know what intelligence is, at least natural human intelligence, that we have reason to believe that we are on the cusp of creating it? Would we not first need to be clear on *that* before we could conclude that we have been doubled, perfected, and replaced? There is no more consensus on the nature of intelligence than there is on morality, either among philosophers or psychologists, but to the contrary a long and ongoing debate that is as old as the first Greek philosophers and as recent as Nagel's 2012 *Mind and Cosmos*.

The following essay is intended as only a first step in staking out the terrain to be discussed. I will not have the opportunity here to develop the three distinctions necessary to an intelligent debate about artificial intelligence: first, the distinction between natural intelligence (NI), common at least to all the higher animals, and natural human intelligence (NHI), unique to us; then the distinction between ANI, artificial narrow intelligence, which presumably we have already invented, and AGI; and finally, and most crucially the distinction between NHI and AGI. This will not be easy or without controversy, on the contrary, we should expect that in the course of seeking clarity on these distinctions, we will have to re-animate historical philosophical debates, between nominalist and realists, for example, between idealists and materialists, and indeed, among monotheists, pantheists and atheists. That things will become messy is no reason to avoid the work. Without this effort, we have no hope of advancing the current debate beyond the materialist prejudices and theological platitudes which currently dog both sides of it.

Everything that I say in what follows will need to be much more carefully developed in the subsequent endeavors of the Working Group on Natural and Artificial Intelligence (WGI), founded at the conference on 'Natürliche und Künstliche Intelligenz im Anthropozän' held 01.–03.03.2020 in Ladenburg, Germany. This preliminary effort is written in anticipation of the larger, collaborative, interdisciplinary work ahead of us. It is a programmatic essay that has as its aim outlining the basic terms to be defined and the arguments to be developed in what might be the most important debate of our time. Without trying to answer all of the questions raised here, it seems clear enough to me that we have a problem. We are trying to build artificial general intelligence without understanding what natural intelligence is. It was this conundrum which led Uwe Voigt and myself to propose the establishment of the WGI, which would draw on the most significant contributions in the philos-

ophy of mind, phenomenology, consciousness studies, cognitive science, theology, and psychology, from the whole history of the Western canon (starting with Aristotle's *De Anima* and extending to contemporary panpsychism debates), to produce an adequate description of the basic features of what makes human intelligence human, and what are the arguments for affirming or denying its existence in non-humans, animal or mechanical. The concrete deliverable would be something like a report on intelligence, for computer engineers and policy makers, which would summarize in layperson's terms what central currents in the Western tradition have meant and still mean by the terms 'intelligence,' 'understanding,' 'rationality,' 'consciousness,' and 'soul'.

2. What is at Issue in the Question Concerning AGI

An ambiguity pervades the current discussion about AGI, an ambiguity about the aim of the project from the beginning. Are we seeking to design a machine that can do all that we do, and more, better than we do it, however it does it? Or are we seeking to design a machine that does what we do *in the way we do it*, i.e., a machine that is not only empirically conscious (cognitively responsive to sense data) but also intelligently and rationally conscious³? And are these two aims separable⁴? For our purposes, it is the second of these two alternatives that is of most interest. The singularity will not be brought about because of the efficiency of our machines at organising the ends we set for them, but because of the capacity of our machines to set ends we have not decided upon: not only machine learning acquiring the

³ The distinction between three levels of consciousness, empirical, intelligent, and rational is drawn from the Canadian Thomist theorist, Bernard Lonergan. Lonergan's immense output is not widely enough known outside of theological circles. As it has as its aim a modern, realist theory of human cognition that can confirm what is true about the Greco-Latin tradition, while developing it in the light of modern probability theory and historical consciousness, it is of direct relevance to the research of the WGI. On the three levels of consciousness, see chapter 11 of his magnum opus *Insight*, 'Self-Affirmation of the Knower'; cf. Lonergan 1957, pp. 343–371.

⁴ See Fjelland (2020), p. 2: "It is possible to pursue this goal without assuming that machine intelligence is identical to human intelligence. For example, one of the pioneers in the field, Marvin Minsky, defined AI as: '[...] the science of making machines do things that would require intelligence if done by men' (quoted from Bolter, 1986, p. 193)."

capacity for intentional thought, which we share with the higher animals, but above all acquiring the capacity for judgment and decision.

In a recent article, Ragnar Fjelland (2020), has looked at the evidence in support of the widespread claim of some computer engineers that we are only decades away from AGI⁵, and concluded that it is exaggerated. Neither algorithmic AI (the brain behind Amazon, YouTube, and countless other consumer service providers), nor more recent advances in creating artificial neural networks, have come close to the promises of AGI. Rather we are producing variations on what Fjelland calls ANI (Artificial Narrow Intelligence): machines that can achieve amazing feats such as *Deep Blue* which beat the world chess champion, Garri Kasparov, in 1997, or *AlphaGo* which defeated the world Go champion, Le Sedol, in 2016, because they were programmed to do nothing else. But humans are good at many things. Specialisation, as anyone who has persisted through a PhD program knows, is a limiting and constraining of natural human intelligence. ‘The overestimation of technology is closely connected with the under-estimation of humans’⁶. What AGI researchers are running up against is the natural ability of ordinary humans to do many things more or less well, even though they cannot explain how it is they do them, and on the basis of this limited and unthematised knowledge, their related ability to understand people very different from themselves and to continue learning. This requires the ‘tacit knowledge’ Michael Polanyi defined as that human ability to learn something complicated from other human, like swimming or riding a bicycle, without having the faintest idea of exactly how one does it (Polanyi 1966). Tacit knowledge, Fjelland argues, has to do with being embodied and inhabiting a world⁷. He concludes that ‘Hubert Dreyfus’s arguments against general AI are still valid’⁸—even some fifty years later! This is because so called ‘general intelligence’ depends upon ‘being-in-the-world’ in Heidegger’s sense of the term, ‘world’ – the ‘referential totality’ [*Verweisungsganze*] which is never an object of knowledge or concern but the horizon of all concerns, plans, tool usages,

⁵ See the survey conducted in 2012/13 by Vincent Müller and Nick Bostrom. The consensus among experts is that there is a fifty per cent likelihood that AI will exceed human intelligence by the middle of this century; cf. Müller/Bostrom (2016). Thanks to Dr. Greg Naterer for drawing my attention to this.

⁶ Fjelland (2020), p. 3.

⁷ Ibid., p. 6.

⁸ Ibid., p. 8.

and knowable objects⁹. Only the existential embodiment, enculturation, and historicity characteristic of *Dasein* (the mode of being unique to the human being) grants one the capacity to perform the countless tasks and to quickly learn countless others, which we take for granted in ordinary life.

I would like to speak in this paper about a different feature of human being that continues to elude AI researchers, and which is under-appreciated in the Heideggerian account: rationality. Heidegger is involved in a not-so-subtle polemic with the whole Western tradition on its ostensible privileging of the theoretical attitude over practical life. This is not at all irrelevant to the question of the nature of human intelligence, of course, as Dreyfus had made plain. But here I do not want to by-pass human theoretical or objective knowing or cognition, but rather to focus on it for a moment. Human cognition I take to be expressed not in the rule-following or in the exhaustive mapping of probabilities, which computers already exceed us at, but in judgments of facts and decisions about what ought to be done in a particular situation. A first obstacle to be removed in the discussion about whether or not AGI in the strong sense of reduplicating NHI is possible is a persistent impoverished understanding of what we doing when we know anything at all.

Reductionist theories of mind seem to abound in AI circles. Reductionism is hardly a new problem in the history of philosophy. Recall Socrates explaining his early enthusiasm for Greek materialism and his disappointment at discovering that it left the one thing most in need of explanation unexplained, the nature of mind (*Phaedo* 95a–100a). He read with interest Anaxagoras's claim that 'it is mind that produces order and is the cause of everything' (97c). He took this to mean that everything was arranged in the way that it was best for it to be, i.e., in Aristotle's terms, that things are ordered according to final causes. Any sound and valid explanation would articulate the final cause of the *explanandum* and make it clear why it was the way that it was. Anaxagoras, however, quickly disappointed Socrates by substituting necessary, physical conditions for the existence of mind for sufficient explanations (the recurring eliminative materialist error). Despite a promising start, Anaxagoras proves himself a reductionist:

It was a wonderful hope, my friend, but it was quickly dashed. As I read on
I discovered that the fellow made no use of mind and assigned to it no cau-

⁹ Heidegger (1927): pp. 111/145.

salinity for the order of the world, but adduced causes like air and aether and water and many other absurdities. It seemed to me that he was just about as inconsistent as if someone were to say, the cause of everything that Socrates does is mind—and then, in trying to account for my several actions, said first that the reason why I am lying here now is that my body is composed of bones and sinews, and that the bones are rigid and separated at the joints, but the sinews are capable of contraction and relaxation, and form an envelope for the bones with the help of the flesh and skin, the latter holding all together, and since the bones move freely in their joints the sinews by relaxing and contracting enable me somehow to bend my limbs, and that is the cause of my sitting here in a bent position. Or again, if he tried to account in the same way for my conversing with you, adducing causes such as sound and air and hearing and a thousand others, and never troubled to mention the real reasons (Phaedo 98e, trans. Hugh Tredennick).

The reductionist—in the 4th century BC or the 21st century AD—purports to explain the whole in terms of the part. Socrates heads off the error in its inception, and Western thought is in the mainstream free of it until late medieval nominalism appears. Now, or at least until recently, reductionism *is* mainstream, particularly in the philosophy of mind. Equipped with colourful neuroimaging, we are repeatedly assuming that a necessary condition without which mind cannot perhaps exist, such as the brain, or the nervous system, is also the sufficient condition for its existence¹⁰.

In the early days of AI debate, philosophers such as John Searle, among the analysts, and Dreyfus, among the continentalists, tried to show the fallacy involved in the assumption that reproducing and improving on the human capacity to manage information would also reproduce human consciousness¹¹. While much has happened in computer science since then, not so much, it seems, has happened in the philosophy of mind. Markus Gabriel is busy popularising neglected ar-

¹⁰ For a fresh take on how to use neuroimaging in a non-reductionist philosophy of mind, see Thompson (2014). By using brain scans to make sense, of all things, classical Indian idealism, Thompson shows that neuroimaging can offer evidence for a theory of mind, but cannot itself serve as the ground for a theory of what mind is.

¹¹ See Searle (1980); Dreyfus (1972). See also Nagel (1974). For a more recent critique of the naive assumptions of AGI, see Smith (2019); Fjelland (2020).

guments culled from the dusty tomes of the German Idealists to refute the new materialists¹². He has good reason to do so: nothing was more evident to Fichte, Schelling, and Hegel, than the irreducibility of mind to its material conditions of operation¹³. David Chalmers's much discussed zombie argument repeats in some ways Searle's Chinese room experiment of the early 80s: a functionalist account of the human difference, which presumes that a machine that passes the Turing test because it acts and responds to questions as humans act and respond, leaves out the very thing in need of explanation, what Chalmers calls 'the hard problem of consciousness', i.e., the question why is there subjective experience in the first place¹⁴?

The question raised by Dreyfus, Nagel and Searle in the 70s and 80s was the following: Is a human intelligence essentially an information processor? If it is, then we have been already supplanted. My cell phone is a much more efficient processor than my brain, which habitually forgets, misjudges, and sometimes deliberately distorts information—even to itself—for various obscure reasons.

¹² See Gabriel (2017).

¹³ See Fichte (1794); Schelling (1797); Hegel (1845), pp. 1–24.

¹⁴ See Chalmers (1995); id. (1996); Searle (1980). Searle's Chinese room argument was intended to show that one could not infer rational consciousness in a machine on the basis of its capacity to correctly respond to questions. A man locked in a room with sufficient time could learn to respond correctly to a series of written questions slipped under the door by a Chinese speaker outside the room—without being able to speak or understand Chinese. It would simply be a matter of learning to produce the signs that were expected; knowledge of what those signs meant would not be necessary. We need not invent such complicated thought experiments to make the point. A child learning his or her multiplication tables by memory is doing the same thing as the man who speaks no Chinese correctly communicating with Chinese symbols. The skilful, publicly validated use of signs does not require insight into meaning—a point to be developed below. There is no intellectual act of understanding (*intelligere*) in the Chinese room experiment or in the memorising of multiplication tables. The later Wittgenstein endeavoured to show that all so-called understanding is nothing but learning the rules for publicly manipulating signs, an argument that is no longer as popular as it once was, but which still needs to be examined in so far as it contests the point I wish to make here: that human understanding is the cognitive, and so immaterial, grasp of sense by mediation of a material sign. See Wittgenstein on 'following a rule'; cf. Wittgenstein 1958, §§ 143–171. See especially § 154: "Try not to think of understanding as a "mental process" at all.—For that is the expression which confuses you. But ask yourself: in what sort of case, in what kind of circumstances, do we say, "Now I know how to go on" [...]" This behaviouristic account of mind is precisely what Nagel (2012) seeks to refute in *Mind and Cosmos*.

But if NHI is not an information processor, then we need to re-open the question of how to best characterise it¹⁵. This is the essential question that must still be addressed as we move forward into the era of machine learning. Like any good question it can be broken down into other, smaller questions. For example, information processing requires the manipulation of signs. At the basic level, every piece of data in a computer can be expressed as some combination of two signs: 0 and 1. But are there other ways of using signs, perhaps more distinctively human, which are not primarily manipulative and pragmatic? Do all animals use signs as stand-ins for objects over which they seek control? Do some animals, human animals most notably, not use signs not only or even primarily as indexicals to facilitate practical activity but also as symbols in a stricter sense of the term, i.e., as mediators of meaning¹⁶? This is the question that I asked in *Thinking Nature* (2019). Drawing on a minority consensus in 20th century theory, with a diversity of representatives in psychology (Jung), the philosophy of science (Cassirer), theology (Tillich) and hermeneutics (Ricœur), I argued for a functional distinction between signs and symbols as key to understanding ‘the human difference’: all symbols are signs but not all signs are symbols. The symbol has a non-indexical function in certain distinctively human forms of discourse. In *Thinking Nature* my concern was the distinction of NHI from NI. In this essay I wish to look at the distinction in terms of the difference between NHI and AI.

To this end I would like to add the following consideration to the question concerning the human difference. What role does the human being’s always marginalized aesthetic capacities play in NHI? The one thing most paleontologists can agree on is that when the modern human appears on earth some

¹⁵ This also raises the question concerning NI and the characterisation of animal consciousness. Is animal consciousness properly characterised as a sign-mediated information processor? I do not have the space to enter into this discussion here, but the question must nonetheless be asked. In this regard see Soentgen on the role of emotion in the inner lives of animals (2018).

¹⁶ We cannot rule out the possibility of forms of NI that are still higher than us, as Thomas Aquinas believed existed. See Aquinas on the reason for positing angelic consciousness, above the human but below the divine (notably for the sake of heeding the principle of plenitude): *Summa Theologica* 1a, q. 50, a. 1. Uwe Voigt has also raised the possibility of higher forms of trans-human intelligence with his theory of the technosphere as a ‘hyper-subject.’ See Voigt (2018).

200,000 years ago, art is left behind, in shattered figurines around their fire pits, and on the walls of caves where they took shelter from the ice age. Is the aesthetic sensibility that makes us so unique among the animal kingdom not more distinctive of our kind of intelligence than the speed with which we solve problems¹⁷?

Before we can be clear that we have created artificial intelligence, we need to be clear on what natural intelligence is, and how widely it is distributed among the earth community, and this clarification, the taxonomy of NI, shall be one of the more important tasks of the WGI. By and large the historical discussion of the nature of mind has neglected this issue and focussed often exclusively on human intelligence or (NHI)¹⁸. A brief review of the discussion concerning NHI in late modern philosophy¹⁹ reveals a focus on three essential marks of *rational* intelligence. Any-

¹⁷ This raises the vexed question (but it cannot be avoided), What is art? What evolutionary purpose, if any, does it serve? Cynthia Freeland describes art as human activity that cannot be reduced to biological aims; cf. Freeland (2001). Palaeolithic art was initially believed to be an instance of 'sympathetic magic,' a ritual using symbols for things over which influence was sought. Such an explanation of early art fit in well with the neo-Darwinian account of human origins, according to which, everything distinctively human emerged in the brain of the ape because it gave the human a natural advantage over other apes. Along this reductionist line, the cave dweller was painting animals in order to secure the success of the hunt. This argument, which I will discuss in more detail below, has since been challenged by who note that Palaeolithic art just as likely had a ritual purpose which had nothing to do with a successful hunt. According to Susanne K. Langer, the Palaeolithic artist was indeed doing ritual magic, but magic is primarily expressive, not pragmatic. See Langer (1957), p. 49: 'Whatever purpose magical practice may serve, its direct motivation is the desire to symbolise great conceptions' However we answer the question concerning the purpose of art, it is clear that the Palaeolithic artist, not unlike the medieval artist, or the contemporary street artist, was expressing the symbols that made manifest the collective identity of his or her people; he or she was making something visible not only as a means to some end, e.g., a successful hunt, but also as an end in itself, and offering the symbols to the community for contemplation, both of its world and itself.

¹⁸ Exceptions include Hegel (1845: 29–152; the 'anthropology'); and the largely forgotten 'psychophysics' of Gustav Fechner (Fechner 1848; 1851; 1860).

¹⁹ Among moderns, in addition to Lonergan, see Peirce (1895) and Husserl (1929). The literature on the nature of mind is indeed vast and will bring us back, as it should, to Aristotle, via his interpreters, in reverse order, Lonergan, Hegel, Scotus, Aquinas, Averroes, Al-Farabi, Avicenna, Plotinus. If Whitehead is correct, that all of Western philosophy is a series of footnotes to Plato, we might also say that all of Western philosophy of mind is a series of footnotes to Aristotle.

thing lacking the capacity for all three cannot be considered intelligent in a human way, or in more precise terms, *rationally* conscious:

1. intentionality
2. rational judgment, including aesthetic judgment
3. moral decision

It would seem that we should attribute some rudimentary form of the first of the three traits to the higher animals, and locate the human difference in the last two. Nothing is more intentional than my cat watching a mouse. Everything about the quality of his attention declares 'aboutness' or 'directedness.' But by the same token, nothing my cat does would justify me in attributing judgment or decision to him.

Missing from the list of essential marks of properly human consciousness is the concept of 'care' or interested and embodied intelligence. It is not clear to me whether this Heideggerian concept, which Dreyfus deployed to refute the very idea of artificial intelligence at MIT in the 70s, is a fourth feature of rational consciousness, or a phenomenologically refined, 'fore-theoretical' interpretation of the three. Care, which Heidegger defines as 'ahead-of-itself-Being-already-in (the world) as Being-alongside entities which we encounter (within-the-world)'²⁰ is a constitutive feature of human being, according to Dreyfus, more essential to us than the capacity to solve problems or process information, and presupposes features, or in Heideggerian language 'existentials' machines manifestly lack, notably the embodiment, enculturation ('thrownness'), and historicity already mentioned. For a machine to be intelligent in a human way, it would have to care about its being, which means it would have been gripped by a troubled history with its being; it would have to be interested in its possibilities for being, and indeed anxious about its death. Care indicates the existential limitations of human being-in-the-world, its thrownness into being, and its call to take up as ground of its being a ground which it did not lay. It presupposes an environment natural to a human existence, i.e., a world. A machine that cares would be a form of being-in-the-world, like us, not a super intelligence or an abstract bodiless mind.

²⁰ Cf. Heidegger (1927), pp.192/237; Dreyfus (1972).

3. Symbolic Thinking as Presupposition of Rational Consciousness

I would also add that a machine that cares would be a machine that inhabits a world mediated by meaning, i.e., it would be a machine capable not only of sign usage but also of symbolic thought. In *Thinking Nature*, I drew on Ernst Cassirer and his student, the now mostly forgotten philosopher of mind, Susanne K. Langer (an important influence on Lonergan's cognitional theory), to make the case that the human difference consists in the special way that the human animal uses signs, as symbols and not merely indices²¹. This was not to revive the tired argument that the human difference is *just* language, for clearly other creatures communicate with signs. My cat meows loudly at noon because he knows that it is time for food. My fifteen-year-old son asks, 'What's for dinner?' every night at 6:30 pm on the same, basically animal impulse, and uses signs, in his case, words, analogously to the way my cat uses its meow. The claim in *Thinking Nature* was first of all to refute the prejudice that humans alone are communicative or sign users: animals which plainly use signs are also to that degree conscious and intentional. Nevertheless, there is a distinctive way that humans use signs, which is at the very core of human culture. If all the higher animals, and perhaps all animals, use signs to communicate with one another, only humans use signs to *express meaning*, i.e., only humans use signs as *symbols*—so I argued. With Langer, I follow Cassirer and draw a sharp distinction between signification, which is a direct indexical reference to a present object or state of affairs, and symbolization, which is an indirect reference to an object in absentia via a showing of meaning. Symbolisation is not confined to language but is also pre-eminently at play in

²¹ See McGrath (2019), pp. 21–25, 87–95; id. (2018). Peirce distinguishes signs into three categories: icons, indices, and symbols; cf. Peirce (1895), p. 13. On the difference between the indexical sign and the symbolic sign see especially Langer (1957), p. 30: 'Man, unlike all other animals, uses "signs" not only to indicate things, but also to represent them'; *ibid.*, pp. 60–61: 'Symbols are not proxy for their objects, but are vehicles for the conception of objects. To conceive a thing or a situation is not the same thing as to 'react toward it' overtly, or to be aware of its presence. In talking about things, we have conceptions of them, not the things themselves; and it is the conceptions, not the things, that symbols directly 'mean'. Cf. Cassirer (1944/1962), pp. 23–41; Lonergan (1971), pp. 57ff. Drawing on German Romantic philosophy of language (Hamann, Herder and Humboldt), Charles Taylor has also argued for the non-indexical quality of verbal symbols. See Taylor (2016), pp. 17–21.

ritual and in art. In fact, most of what we do in language is not signifying in the way that the meowing cat can be said to be signifying his hunger²². A meaning or sense is often (though not always) evoked by a symbol for the sake of consideration, and not merely as a means to an end. When I symbolize something by means of its associated senses—and connotation is for the most part not univocal but metaphorical and analogical, for symbols are most alive in ambiguity²³—I am not seeking to achieve any practical aim in the world, or to evoke a response from the hearer, as I do when I call out someone's name. Rather I symbolize for the sake of contemplative consideration, or to use the ancient Greek term, *theorein*. Such forms of communication are examples of what Aristotle calls *theoria*, attending to an intelligible essence for the sake of knowing it²⁴. On this line, Aristotle's *zoon logon echon*, or Cassirer's *animal symbolicum* (i.e., 'thinking nature,' i.e., not only the nature that is thought but the nature that thinks itself)—human being—is first and foremost contemplative being. Once we have attended to our practical needs—communicatively collaborating with one another for the sake of securing food, shelter, and sexual partners—we have the leisure requisite for contemplating the sense of the things that are. This can happen in a religious way, when I attend a celebration of the Eucharist at my parish church. It can also happen in a high-brow way, when I visit a gallery to look at fine art. But much more commonly, it happens in a low-brow, quotidian way, when, for example, I engage in idle gossip with my partner over breakfast or watch the news after dinner. In each of these

²² See Langer (1957), p. 31: 'Most of our words are not signs in the sense of signals. They are used to talk about things, not to direct our eyes and ears and noses toward them.' Humans not only, or even primarily, signify things with verbal signs, they denote things by connoting meanings through verbal symbols. In Langer's terms, a symbol 'denotes' its referent or signified object, via a 'connoting' of its sense or senses. By insisting on four terms in symbolization—sign, denoted referent, connoted meaning, and object—Langer breaks with the structuralism that eventually won the day. Structuralism recognises only two terms in a symbolic structure, the signifier, which is an arbitrary sign, and the signified, which is a concept, with no direct relation to the real, but which is only determined negatively by its differential relation to other concepts. Structuralism is the apogee of nominalism and severs the relation of the symbolic to the real. See Saussure (1916); Lacan (1966). The hegemony of structuralism over continental thought in the 20th century is no doubt one of the reasons Langer's work has been forgotten. In addition to Philosophy in a New Key, see Langer (1967); id. (1972).

²³ Ricoeur (1970), pp. 3–19; Tillich (1999), pp. 42–48.

²⁴ De an. 412a23; 417a28; Eth. Nic. 1146b33; 1177a18; Metaph. 1048a34; 1072b24; 1087a20.

instances, religious, aesthetic, quotidian, I am doing something other animals apparently do not do (at least there is no evidence for it).

The human contemplative enjoyment of meaning seems to be older than civilisation. One of the things that distinguishes the remains of the fires around which early humans assembled from the remains of the fires made by their contemporary Neanderthals is that human fires were much deeper and more established, by distinction from the Neanderthal fires which were made quickly, as need required, and abandoned as soon as they were no longer needed. Human fires were in fact, hearths, around which the human tribe lingered after cooking and eating, and to which they returned, year after year, leading some paleontologists to hypothesize that such lingering led naturally to ritual activities, myth making, or even simply casual conversation, i.e., the more sophisticated usage of signs as symbols which gave rise to the higher intelligence of this species descended, among other species such as *Homo neanderthalensis*, from a common ancestor, *Homo erectus*²⁵.

One other example to make it clear that we are not speaking only or even primarily about language: the oldest piece of art in the world is the Hohlenstein-Stadel Löwenmensch, a prehistoric ivory sculpture, 31.1 cm tall and 5.6 cm wide, of a female humanoid figure with the head of a lion. Carbon dated to between 35,000 to 40,000 years ago, the Löwenmensch precedes the cave paintings of Lascaux by some 20,000 years. It was made by people who hunted the huge mammals that grazed along the edge of the retreating glaciers in Europe during the last ice age, and sheltered in caves from the other mammals that preyed upon them. Palaeontologists who re-enacted the production of such a piece of art, making use of the kinds of stone tools available to those who carved the Löwenmensch, found that it took more than 370 hours of delicate, highly skilled work, to complete the task²⁶. Asked why a tribe of humans struggling to stay alive in the last ice age would have allowed one of their members to be exempt from subsistence work to create art to this extent, Jill Cook, Curator of Palaeolithic collections at the British Museum, answered, it was to have one among them express ‘a relationship to things unseen, to the vital forces of nature’ (Cook in MacGregor 2017). Neo-Darwinians will argue that this is an example of art developing as a form of sympathetic magic on the questionable assumption that every human ability must be explained in terms of evolutionary advantage.

²⁵ Wynn/Coolidge (2011), pp. 112ff.

²⁶ Cook (2013).

The palaeolithic artist and his or her tribal patrons, on the neo-Darwinian line, were trying to control their dangerous environment. Ostensibly for the same reason that palaeolithic artists developed the skills needed to produce the exquisite paintings of the Lascaux caves, our Cro-Magnon fore-bearers are assumed to have been simply trying to get an edge on the large mammals competing with them for survival. It is more reasonable to assume that ice age artists were doing the same thing we do when we make art, or make it possible for some of us to develop the skills needed to do so, by subsidising the lives of artists with grants and scholarships: they were, in Langer's language, 'symbolically transforming' their common experience and so elevating signs, and their minds which depend on them, from the practical and in-



Fig. 1: The Löwenmensch. Prehistoric ivory sculpture. The oldest known uncontested example of figurative art. Carbon dated to 40,000 BC. Discovered in the Hohlenstein-Stadel cave in 1939. (Source of image: Wikimedia Commons).

dexical into the symbolic and the domain of meaning. They were using signs as symbols for the sake of contemplating the meaning of their day-to-day reality, and they were doing it for no other reason than that it pleased them to do so. By contemplating the metaphysical in the shape of the Löwenmensch, they were also contemplating themselves, for to think anything symbolically or contemplatively is to also think the thinking that thinks the thing. Indeed, as phenomenologists have been arguing for a century, we only think ourselves thinking by thinking about something²⁷.

Symbolic thought, by distinction from significative thought, is the condition for the possibility of *rational* consciousness. Consciousness need not be rational, as we see from its instantiation in other animals and in ourselves some of the time; it is often nothing more than a complex response to sensation, and so continuous with the stimulus response found in the simplest living organisms, in plants as well as simple animals. The human difference is something beyond sensitive or ‘empirical consciousness’²⁸. It consists not only in the awareness of sensitive experience and the capacity to imaginatively respond to it, but in the capacity to *transcend* our subjectivity and inquire into, and to some degree understand, the nature of that which we experience.

4. Revisiting the Argument against Functionalism

This capacity for symbolically-mediated objectivity has been repeatedly invoked by philosophers of mind to refute the so-called functionalist argument. Rooted in Alan Turing’s test of the same name, which was designed to prove the indiscernibility of a sufficiently sophisticated mechanical response to a question from a human re-

²⁷ See Husserl (1929), pp. 33–37; Lonergan (1957), pp. 344–346. Cf. Aristotle, *Metaph.* 1072b20–25: ‘Thought thinks on itself because it shares the nature of the object of thought: for it becomes an object of thought in coming into contact with and thinking its objects, so that thought and the object of thought are the same. For that which is capable of receiving the object of thought, i.e. the essence, is thought. But it is active when it possesses this object. Therefore the possession rather than the receptivity is the divine element which thought seems to contain, and the act of contemplation is what is most pleasant and best.’

²⁸ Lonergan (1957), p. 346.

sponse, and the later Wittgenstein's behaviourism, the functionalist argument holds that for a machine to be considered intelligent it is enough for it to respond and act in the same outwardly observable fashion that a human being responds. The counter argument holds that a generally intelligent machine would need to not only do what humans do, but also do it in the *way* humans do it. It would need to act for *reasons*, i.e., its acts would need to be the outcome of judgments and decisions, i.e., the result of a reasoning process, which is oriented to structures of intelligibility that are not reducible to our thinking them. One can memorize mathematical formulae without understanding them. And when one thereby 'solves' math problems, one acts in the same way that a machine responds to input on a keyboard. The machine does not understand that $2+3=5$; it responds to the input in the way it is determined to respond. An elephant can be trained to use a paint brush and produce abstract pictures that can be sold for a good price on the art market²⁹. But no one seriously believes that the elephant is making art for the same reasons that the human being makes art. Rational consciousness appears to require more than the capacity to respond to stimuli; it appears to be more than a mechanical reaction: it judges states of affairs and whenever it does so correctly, it reaches beyond the circumstances and the practical need of the judge. To judge rationally, whether of a matter of fact or of concern, is to transcend need and circumstance and affirm or deny the truth of what is at issue. How exactly humans do this, and why they should have evolved in such a way as to be able to do it, is the theme of Thomas Nagel's *Mind and Cosmos*.

According to Nagel, a reductionist theory of evolution, which would explain mind in terms of the evolution of material processes, and so all animal behaviour in terms of naturally selected advantages, cannot make sense of rational judgment. One way the reductionist deals with this problem is by denying the existence of mind altogether. Nagel notes that the denial of the existence of mental states was the strategy of 20th century behaviourist philosophers of mind, such as Ryle and Wittgenstein.

The names of mental states and processes were said not to be referring expressions. Instead, mental concepts were explained in terms of their observable behavioural conditions of application—behavioural criteria or ascertainability conditions rather than behavioural truth conditions³⁰.

²⁹ See <https://www.youtube.com/watch?v=gjOydUjjDos>, retrieved 17.07.2021.

³⁰ Nagel (2012), p. 38.

The problem with these arguments according to Nagel is that they leave out exactly that which is to be explained, the first-person experience of being a mind, ‘what it is like’ to be conscious of something,

the way sugar tastes to you or the way red looks or anger feels, each of which seems to be something more than the behavioural responses and discriminatory capacities that these experiences explain³¹.

Assuming that denying the existence of mental states and reducing understanding to observable rule following not an adequate strategy for addressing the hard problem of consciousness, Nagel concludes that ‘conscious subjects and their mental lives are inescapable components of reality not describable by the physical sciences’³². Along a similar line of argumentation, mental states cannot be held to be identical to the brain states that underlie them. It is conceivable that there could be brain states without any mental states (the hypothesis of Chalmers’s zombie argument)³³. Therefore, if there is something called a mental state, it is not identical to a state of the brain or any other material configuration for that matter (for example, the circuitry of a computer); as such it cannot be explained as only a product of material evolution.

Nagel’s main argument, going in exactly the opposite direction of Dreyfus’s defence of the human difference on the basis of ‘care,’ zeroes in on the objectivity of judgment, whether epistemic judgments, concerning an objective state of affairs, or moral judgments, concerning right and wrong. Along the materialist neo-Darwinian line, he notes, a judgment can be nothing more than a strategic, self-interested move by an organism trying to get one up on its competitors in evolution. If this were true, then the history of science, and human morality—indeed, all our cultural achievements, from ancient religion to quantum physics and modern art— would be explicable as naturally-selected products of evolution. The capacities for science and art would only have developed in us because they gave us an evolutionary advantage. There would therefore be no sense in speaking about objectivity or truth, then or now, for evolutionary determinism is still driving our

³¹ Ibid.

³² Ibid., p. 41.

³³ See Chalmers (1996), p. 96.

minds. As the most recently randomly selected bundle of animal attributes, we contemporary humans only call something true or false because it is in our interest to so call it³⁴. But this would mean that the theory of evolution itself is held to be true, not because it offers us the more coherent and adequate account of the facts of geological time, but because it is in our interest to affirm it as true. Should creationism prove more advantageous (and for a sizeable minority, the jury is still out on this), then creationism shall be justified as true over evolution. A theory is not in our interest because it corresponds to the fact, but because thinking it so gets us one up. Plainly, however, the intention of the scientist who insists on the truth of evolution against his objector, for example, Richard Dawkins debating Rowan Williams at the University of Oxford in 2012, is not to advance his thesis because he believes it to be more advantageous to believe it (although he might also think that), but because he believes it to be true, and the other thesis to be false³⁵. For reasons such as this Nagel argues that any theory of evolution which purports to explain the mental in terms of the physical and to reduce the human difference to a naturally selected evolutionary advantage, commits the ‘functionalist’ fallacy. It collapses the reasons for a judgment into the outwardly performed act of judging. We no doubt developed the capacity to reason in the course of evolution, but reason itself is not a mere expression of natural self-interest. ‘Merely to identify a cause is not to provide a significant explanation, without some understanding of why the cause produces the effect,’ Nagel writes, in effect repeating Socrates’ objection to Anaxagoras³⁶.

Of most interest for the purposes of this paper is Nagel’s distinction between consciousness and reason³⁷. Consciousness in its simplest form might be merely sophisticated stimulus response and so explicable as having evolved because of the natural advantage it gives certain kinds of life over others. But rational intelligence does not merely self-interestedly react to stimuli; rather, it *disinterestedly* responds to objective truth and value. Indeed, the affirmation of a truth is often not in our

³⁴ It is worth remembering that this evolutionary relativism is the very core of Nietzsche’s perspectivism and historicism. See in particular Nietzsche’s early work, ‘On the Uses and Abuses of History’: Nietzsche (1874).

³⁵ See Nagel (2012), p. 81: ‘Any evolutionary account of the place of reason presupposes reason’s validity and cannot confirm it without circularity.’

³⁶ Ibid., p. 45.

³⁷ Ibid., pp. 71ff.; cf. Lonergan (1957), pp. 346–348.

individual egoic or group-centred interest, as psychoanalysts have been pointing out for a century. One could by extension imagine that some truths are not in our interest as a species. The capacity to intelligently respond to truth with a reasoned judgment about the state of affairs regardless of what the judger would prefer to believe, cannot be solely determined by evolutionary advantage.

Thought and reasoning are correct or incorrect in virtue of something independent of the thinker's beliefs, and even independent of the community of thinkers to which he belongs³⁸.

Nagel is hardly the first to draw the distinction between consciousness and reason, which can be traced back to Aristotle, and in its Aristotelian registers has been most developed by Lonergan as the difference between empirical and intelligent consciousness³⁹. Nor is Nagel the first to use the distinction to refute a reductionistic, materialist account of mind. Few remember that Husserl's phenomenology originated in a debate with what was then called 'psychologism,' the reductionist argument, emerging out of late 19th century positivism, that judgments are nothing more than the effect of certain psychological conditions or events. Psychologism amounted to a denial of the validity of logic in Husserl's view. Logic must be more than a psychological condition determining how we should judge; Husserl insisted, the validity or invalidity of a judgment must be logically independent of the judgment.

Logical laws, taken in and for themselves, are not normative propositions at all in the sense of prescriptions, i.e., propositions which tell us, as part of their *content*, how one *should* judge⁴⁰.

Inspired by Husserl's argument, and especially that of Husserl's star student, Emil Lask, Heidegger wrote his doctoral dissertation defending logic against psychologism⁴¹. For Heidegger, the undeniable and over-ruling sense of logical validity is a

³⁸ Nagel (2012), p. 72.

³⁹ Lonergan (1957), pp. 346–348.

⁴⁰ Husserl (1900): vol. 1, section 4.

⁴¹ See Heidegger (1913); Lask (1911).

phenomenological indication that judgment transcends the psychological conditions that might accompany it⁴².

The key to the distinction between merely sensitive consciousness and rational consciousness is therefore *judgment*. All consciousness is intentional, but not all consciousness is or needs to be judgmental. With judgments, either noetic or evaluative, we enter what Robert Sokolowski calls ‘the space of reasons’⁴³. For a machine to do most of the things we do, it need not possess rational consciousness. But for a machine to supplant us on the planet, it must possess the power, and assume the risk, of judgment. AI may improve on us with regard to calculative ability and efficiency at optimising the conditions of human flourishing, but it will not replace us as the mind of nature, the microcosmic mirror of the whole, so long as it does not possess the capacity to judge and decide. Without symbolic consciousness, which would allow it the distance from its being to make judgments of truth and falsehood, right and wrong, it will be merely a hyper-efficient animal.

5. Conclusion

This, then, is the Holy Grail of AGI research: not only the functional reduplication of the activities which we now associate with NHI, but the design of a machine that will do the things we do *in the same way* that we do them, albeit with much greater efficiency and evolutionary capacity. The aim of strong AGI is nothing less than the mechanical reduplication of the human difference. The most ambitious and speculatively inclined AGI researchers are not assuming a weak sense of consciousness such as might be predicated of all beings capable of responding to stimuli, from the sea urchin to the robot, but a strong sense of consciousness, consciousness as the ca-

⁴² McGrath (2006), pp. 93–97.

⁴³ See Sokolowski (2000), p. 116: ‘To think is not to behave in a certain way or to be determined by certain brain-dependent mental events (which may be necessary to thought but are not sufficient for it). To think is to enter as agents into the space of reasons.’ Cf. Nagel (2012), p. 79: ‘When we rely on systems of measurement to correct perception, or probability calculations to correct intuitive expectations, or moral or prudential reasoning to correct instinctive impulses, we take ourselves to be responding to systematic reasons which in themselves justify our conclusions, and which do not get their authority from their biological origins.’

capacity for objective, rational judgment, for knowledge in the full sense of the term—*theoria*, not just *praxis*, and *poiesis*, not just *techne*: consciousness that can produce imputable judgments, and is capable of speculative and aesthetic contemplation. Regardless of whether or not such a thing proves possible, the aim itself forces philosophy to clarify how rational judgment, decision, and contemplative perception distinguish human consciousness from other forms of consciousness.

Until a machine gives us reasons to think that it has attained symbolic consciousness and that it now, like us, takes a *theoretical* interest in questions of truth and falsehood, that it too is sometimes driven by a disinterested desire to know, i.e., to contemplate the meaning of its existence, either aesthetically, philosophically, or religiously, we shall have no reason to recognise it as rationally intelligent, i.e., intelligent in the human sense of the term. To the question, What would count as evidence?, we can only point to those cultural products which most plainly exhibit our contemplative impulse and capacity for symbolic mediation, i.e., to art, philosophy, and religion. A machine that had become artistically expressive, philosophically perplexed, or religious would indeed be worthy of our recognition as rationally intelligent. Of course it might always be duping us for its own evolutionary advantage. We could never be sure, just as we are never so sure about each other.

Bibliography

- Bostrom, Nick (2015): *What happens when our computers get smarter than we are*, [https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are?language=en, retrieved 19.07.2021]
- Cassirer, Ernst (1944): *An Essay on Man: An Introduction to the Philosophy of Human Culture*, New Haven-London
- Chalmers, David (1995): Facing up to the problem of consciousness, in: *Journal of Consciousness Studies* 2/3, pp. 200–219
- (1996): *The Conscious Mind: In Search of a Fundamental Theory*, Oxford
- Chesterton, Gilbert Keith (1925): *The Everlasting Man*, London
- Cook, Jill (2013): *Ice Age Art: Arrival of the Modern Mind*, London
- Dreyfus, Hubert (1972): *What Computers Can't Do. A Critique of Artificial Reason*, New York
- Fechner, Gustav (1860): *Elemente der Psychophysik*. English translation of volume 1 of 2: *Elements of Psychophysics*, New York 1966

- (1901): *Zend-Avesta oder über die Dinge des Himmels und des Jenseits. Vom Standpunkt der Naturbetrachtung*, Leipzig
- (1908): *Nanna oder über das Seelenleben der Pflanzen*, Leipzig
- Fichte, Johann Gottlieb (1794): *Die Wissenschaftslehre*. English translation: *The Science of Knowledge*. Ed. and trans. Peter Heath and John Lachs, Cambridge, U. K., 1982
- Fjelland, Ragnar (2020): Why General Artificial Intelligence Will Not Be Realized, in: *Humanities and Social Sciences Communications* 7/10, [<https://doi.org/10.1057/s41599-020-0494-4>, accessed June 22, 2021]
- Freeland, Cynthia (2001): *But Is It Art? An Introduction to Art Theory*, Oxford
- Gabriel, Markus (2017): *I Am Not a Brain: Philosophy of Mind for the 21st Century*, Cambridge
- Hegel, Georg Wilhelm Friedrich (1845): *Hegel's Philosophy of Mind. Being Part Three of the Encyclopedia of the Philosophical Sciences (1830). Together with the Zusätze in Boumann's Text*. Trans. A.V. Miller. Ed. William Wallace, Oxford, U.K., 1971
- Heidegger, Martin (1913): *Die Lehre vom Urteil im Psychologismus*, in: id., *Frühe Schriften. 1912–16* (Gesamtausgabe, Band 1, Ed. Friedrich-Wilhelm von Herrmann), Frankfurt am Main 1978
- (1927): *Being and Time*. Trans. John Macquarrie and Edward Robinson, San Francisco 1962
- Husserl, Edmund (1900): *Logical Investigations*, Two volumes. Trans. Dermot Moran, London 2001
- (1929): *Cartesian Meditations: An Introduction to Phenomenology*. Trans. Dorion Cairns, Dordrecht 1999
- Lacan, Jacques (1966): *Écrits. A Selection*. Trans. Bruce Fink, New York 2002
- Langer, Susanne K. (1957): *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, Cambridge, Mass.
- (1967/1972): *Mind: An Essay on Human Feeling*. Volume 1–2, Baltimore, MD
- Lask, Emil (1911): *Die Logik der Philosophie und die Kategorienlehre. Eine Studie über den Herrschaftsbereich der Logischen Form* (Gesammelte Schriften, Band 2, hg. Eugen Herrigel), Tübingen
- Loneragan, Bernard (1957): *Insight: A Study in Human Understanding*. Ed. Frederick E. Crowe and Robert M. Doran, Toronto 1992
- MacGregor, Neil (2017): *The Beginnings of Belief. Living with the Gods*, BBC Radio 4, 23 October
- McGrath, Sean J. (2006): *The Early Heidegger and Medieval Philosophy: Phenomenology for the Godforsaken*, Washington, D.C.
- (2018): In Defense of the Human Difference, in: *Environmental Philosophy* 15/1, pp. 101–115
- (2019): *Thinking Nature: An Essay in Negative Ecology*, Edinburgh

- Müller, Vincent/Bostrom, Nick (2016): 'Future Progress in Artificial Intelligence: A Survey of Expert Opinion', in: Vincent C. Müller, ed., *Fundamental Issues in Artificial Intelligence*, Berlin, pp. 553–571
- Nagel, Thomas (1974): 'What is it Like to be a Bat?', in: *The Philosophical Review* 83/4, pp. 435–450
- (2012): *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*, Oxford
- Nietzsche, Friedrich (1874): 'On the Uses and Abuses of History', in: *Untimely Meditations*. Trans. R. J. Hollingdale, Cambridge 2012
- Schelling, Friedrich Wilhelm Joseph (1797): *Ideen zu einer Philosophie der Natur als Einleitung in das Studium dieser Wissenschaft* (SW II: pp. 1–73). English translation: 'Ideas on a Philosophy of Nature as an Introduction to this Science'. Trans. Priscilla Hayden-Roy, in: *Philosophy of German Idealism*. Ed. Ernst Behler, New York 1987, pp. 167–202
- Sokolowski, Robert (2000): *Introduction to Phenomenology*, Cambridge
- Peirce, Charles Sanders (1895): 'Of Reasoning in General', in: *The Essential Peirce*. Volume 2. Ed. by the Peirce Edition Project, Bloomington, IN, 1998, pp. 11–26
- Polanyi Michael (1966): *The Tacit Dimension*, Chicago 2009
- Ricoeur, Paul (1970): *Freud & Philosophy: An Essay on Interpretation*. Trans. Denis Savage, New Haven-London
- Saussure, Ferdinand de (1916): *A Course in General Linguistics*. Trans. Roy Harris, London 2013
- Searle, John (1980): 'Minds, Brains and Program', in: *Behavioural and Brain Sciences* 3, pp. 417–57
- Smith, Brian Cantwell (2019): *On the Promise of Artificial Intelligence: Reckoning and Judgment*, Cambridge, Mass.
- Soentgen, Jens (2018): *Ökologie der Angst*, Berlin
- Taylor, Charles (2016): *The Language Animal: The Full Shape of the Human Linguistic Capacity*, Cambridge, MA
- Tegmark, Max (2017): *Life 3.0: Being Human in the Age of Artificial Intelligence*, New York
- Thompson, Evan (2014): *Waking, Dreaming: Self and Consciousness in Neuroscience, Meditation, and Philosophy*, New York
- Tillich, Paul (1999): *The Essential Tillich: An Anthology of the Writings of Paul Tillich*. Ed. F. Forrester Church, Chicago
- Voigt, Uwe (2018): 'Inside the Anthropocene', in: *Analecta Hermeneutica* 10, [https://journals.library.mun.ca/ojs/index.php/analecta/article/view/2057/1647, retrieved 19.07.2021]
- Wittgenstein, Ludwig (1953): *Philosophical Investigations*. Trans. G.E.M. Anscombe, Hoboken, NJ
- Wynn, Thomas/Coolidge, Frederick L. (2011): *How to Think Like a Neanderthal*, Oxford

Künstliche Intelligenz im Lichte der Technikphilosophie. Ein Überblick unter besonderer Berücksichtigung des Mensch-Natur-Technik-Verhältnisses

Thomas Heichele

Zusammenfassung

Dieser Aufsatz behandelt aus dem Blickwinkel der Technikphilosophie einige philosophische Herausforderungen der Künstlichen Intelligenz (KI), wobei ein besonderer Fokus auf das Mensch-Natur-Technik-Verhältnis gerichtet wird. In diesem Zusammenhang wird zuerst die Technikphilosophie als sinnvoller Ort der Reflexion über KI ausgewiesen, bevor notwendige Begriffsklärungen vorgenommen werden. Daran schließt sich unter besonderer Berücksichtigung der KI eine Analyse der Revolutionen im Mensch-Natur-Technik-Verhältnis an. Den Abschluss bildet die Anwendung klassischer technikphilosophischer Topoi auf die KI.

Abstract

This paper addresses from the perspective of the philosophy of technology some philosophical challenges of artificial intelligence (AI), with a particular focus on the relationship between human, nature and technology. The first task is to explain why philosophy of technology is a useful place to reflect on AI, before making some necessary clarifications of terms. This is followed, with special reference on AI, by

an analysis of the revolutions in the human-nature-technology relationship. Finally, there is an application of classical topoi of philosophy of technology to AI.

1. (Technik-)Philosophie als sinnvoller Ort der Reflexion über Künstliche Intelligenz

Sowohl die Schaffung von als auch der Umgang mit Künstlicher Intelligenz (KI) verlangt aufgrund der bereits auf den einzelnen Ebenen vorherrschenden und im Zusammenspiel noch einmal verstärkten enormen Komplexität, die über wissenschaftliche Aspekte hinaus u. a. soziale und psychische Faktoren betrifft, nach interdisziplinärer Zusammenarbeit und ist ein multidimensionales Unterfangen, das neben deskriptiven Analysen auch normative Setzungen erfordert¹. Bei der Bewältigung dieser Herausforderungen kann und muss die Philosophie als Orientierungswissenschaft mit einem (mindestens) dreifachen Anspruch² eine Schlüsselrolle spielen³: Als Universalwissenschaft strebt sie ohne Beschränkung auf ein bestimmtes Gegenstandsgebiet nach universalen Einsichten mit Blick auf das Wahre, das Gute und das Schöne, als Reflexionswissenschaft ist sie die „Anwendung von Vernunft auf ihr eigenes Tun“⁴ und reflektiert das menschliche Denken und Handeln (sowie damit auch sich selbst), und als Metawissenschaft umgreift sie die Erkenntnisse der anderen Disziplinen, setzt diese miteinander in Verbindung und beleuchtet die grundsätzlichen Voraussetzungen des wissenschaftlichen Weltzugangs.

Angewandt auf die vielfältigen Aufgaben und Schwierigkeiten, vor die KI uns stellt, bedeutet dies u. a., dass es an der Philosophie ist, eine rationale Orientierung zu ermöglichen und beispielsweise wesentliche Beiträge im Kontext von Begriffsklärungen, ethischen Bewertungen und der Auslotung prinzipieller technischer (Un-)Möglichkeiten zu liefern. Es liegt an ihr, die generelle Bedeutung von KI für den Menschen zu reflektieren, Interdisziplinarität im Rahmen der KI-Forschung sowohl als Ort des Austauschs als auch als treibende Kraft zu fördern und (sowohl

¹ Einen hervorragenden Einstieg in die Vielschichtigkeit notwendiger Reflexionen im Kontext der KI liefert Rosengrün (2021).

² Vgl. explizit zum dreifachen Anspruch der Philosophie Illies (2006), S. 21.

³ Vgl. zu Potenzial, Wesen und Rolle der Philosophie im 21. Jahrhundert Heichele (2020c).

⁴ Rosenberg (2009), S. 18.

wissenschaftsintern als auch wissenschaftsextern für die gesamte Gesellschaft) einzelwissenschaftliche KI-Erkenntnisse hinsichtlich einer möglichst vollständigen und konsistenten Gesamtschau einzuordnen. Mit ihrem mannigfaltigen Potenzial ist die Philosophie für die Erforschung der KI nicht nur eine entscheidende Voraussetzungswissenschaft⁵ und reflektierende Wissenschaft höherer Ordnung, sondern sie hat auch eine zentrale Rolle beim konkreten wissenschaftlichen Fortschritt inne⁶.

Innerhalb der Philosophie gibt es mit der Technikphilosophie eine Disziplin, die in besonderer Weise für die bestehenden und kommenden Aufgaben einer philosophischen Reflexion der Künstlichen Intelligenz prädestiniert ist⁷. Obschon die intensive Beschäftigung mit Technik seit jeher Bestandteil der Philosophie ist, ist die Technik-

⁵ Diese Tatsache wird – wenngleich in philosophisch defizitären Detailausarbeitungen – z. B. auch im derzeitigen KI-Standardwerk schlechthin anerkannt, in dem der Philosophie KI-relevante Kompetenzen bei der Beurteilung der Möglichkeit einer Anwendung von formalen Regeln im Kontext gültiger Schlüsse, bei der Analyse der Beziehung von Gehirn und Geist, bei der Untersuchung des Ursprungs von Wissen und bei der Beantwortung der Frage, wie Wissen zu Aktionen führt, zugeschrieben werden. Vgl. Russell/Norvig (2012), S. 27–29.

⁶ Dieser Punkt wird mitunter sowohl hinsichtlich des wissenschaftlichen Fortschritts im Allgemeinen als auch bezüglich einzelwissenschaftlicher Entwicklungen im Speziellen kontrovers diskutiert. Vgl. jedoch zur generellen Relevanz der Philosophie für die Einzelwissenschaften z. B. Heichele (2020c), S. 12–16 und S. 22f., sowie exemplarisch und ausführlich zur Rolle der Philosophie bei der Entstehung der neuzeitlichen Physik Heichele (2016). Die Rolle der Philosophie beim ideengeschichtlichen und wissenschaftshistorischen Hintergrund der KI – u. a. bei Aristoteles (384–322 v. Chr.) und der Erkenntnis, dass bestimmte Schlüsse unabhängig von ihrem Inhalt allein aufgrund der Form gültig sind (Syllogistik), über Gottfried Wilhelm Leibniz (1646–1716) mit dem Ideal (der maschinellen Verarbeitung) einer symbolischen Universalsprache (*characteristica universalis*) bis hin zu Gottlob Freges (1848–1925) Arbeiten zu (m Zusammenhang von) Logik und Sprachphilosophie – wird z. B. in Berlinski (2000), Buchanan (2005) und Davis (2011) deutlich, die Notwendigkeit der Philosophie für die aktuelle (und zukünftige) KI-Forschung z. B. in Mainzer (2019) und Rosengrün (2021). Es sollte auch nicht übersehen werden, dass Alan Turings (1912–1954) 1950 erschienener Aufsatz „Computing Machinery and Intelligence“, in dem er u. a. sein heute als Turing-Test bekanntes „Imitation Game“ vorstellte und der (auch wegen der dort behandelten Themen des Maschinellen Lernens und der Evolutionären Algorithmen) einen Meilenstein für die moderne KI-Forschung bedeutete, in einer philosophischen Fachzeitschrift veröffentlicht wurde.

⁷ Gute Einführungen in die Technikphilosophie sind z. B. Nordmann (2008), Kornwachs (2013) und Zoglauer (2002b). Einen überaus reichhaltigen und gehaltvollen Fundus technikphilosophischer Reflexionen findet man in Gallee (2003), einen ersten Zugang

philosophie als eigenständige philosophische Subdisziplin erst ein Kind des späten 19. Jahrhunderts⁸: Als ihre „Gründungsschrift“ wird oft⁹ Ernst Kapps (1808–1896) 1877 erschienenes Werk „Grundlinien einer Philosophie der Technik“¹⁰ herangezogen. Nicht zuletzt aufgrund der Tatsache, dass bis heute kein Konsens bezüglich der Explikation des Technikbegriffs besteht¹¹, gibt es keine allgemein anerkannte Grenzziehung hinsichtlich des exakten Gegenstandsbereichs und der genauen Aufgabe der Technikphilosophie. Besonders fruchtbar sind Auffassungen, wonach Technikphilosophie generell die Frage nach dem „Wie“ der Welterschließung behandelt¹² bzw. dass es in ihr allgemein um die Wiederholung der gesamten Bandbreite philosophischer Fragen unter besonderer Berücksichtigung der Technik¹³ geht. Dieser Ansatz erfährt eine tiefere Plausibilität, wenn man ‚Technik‘ – dies ist je nach Kontext z. B. auch bei ‚Natur‘ oder ‚Kultur‘ sinnvoll – als Reflexionsbegriff verwendet, durch den bestimmte Hinsichten ermöglicht werden¹⁴. Gerade diese Kombination aus thematischer Offenheit einerseits und spezifischem Reflexionszugang andererseits macht die Technikphilosophie zu einem idealen – wenngleich natürlich nicht dem einzig sinnvollen – Ort des philosophischen Nachdenkens über KI.

In der Technikphilosophie bestehen – vor dem Hintergrund ihrer Vielfältigkeit und angesichts des spezifischen Reflexionszugangs zu ‚klassischen‘ Themen nicht verwunderlich – besondere (und über die in den philosophischen Subdisziplinen generell anzutreffenden hinausgehende) Querverbindungen zu anderen philosophischen Subdisziplinen. Dies betrifft u. a. die Anthropologie beim Blick auf den Menschen als herstellendes und Technik verwendendes Wesen und die

zu typischen Problemen und Anwendungen der Technikphilosophie z. B. in Heichele (2020a).

⁸ Vgl. zur Geschichte der Technikphilosophie z. B. Heichele (2016), Kap. 2.1 und 2.2.

⁹ Vgl. z. B. Hubig (2006), S. 17, Nordmann (2008), S. 9, Ropohl (2009), S. 13 und Zoglauer (2002a), S. 9.

¹⁰ Kapp (1978).

¹¹ Ein erster Überblick über die verschiedenen gebräuchlichen Technikbegriffe und die Schwierigkeiten einer allgemeinen Grenzziehung findet sich z. B. in Heichele (2016), S. 18–20.

¹² Vgl. z. B. Hubig (2006), S. 25.

¹³ Vgl. z. B. Nordmann (2008), S. 10–19.

¹⁴ Vgl. zu ‚Technik‘ als Reflexionsbegriff z. B. Grunwald/Julliard (2005), Hubig (2006), S. 229–234 und Nordmann (2008), S. 12f. Eine diesbezüglich kritische Haltung findet sich z. B. in Ropohl (2010), Kap. 3.3.

Ethik im Rahmen der Analyse von Kriterien und Konsequenzen technischen Handelns. Ein starker Bezug zu Erkenntnis- und Wissenschaftstheorie entsteht beispielsweise bei der Beleuchtung des Zusammenhangs von Technik und Wissen(-schaft), Metaphysik und Naturphilosophie spielen exemplarisch bei der Frage nach dem Wesen der Technik bzw. bei der Bestimmung des Natur-Technik-Verhältnisses eine wesentliche Rolle. Diese Querverbindungen (und noch viele mehr) treten auch bei der technikphilosophischen Beschäftigung mit KI zutage, wenn es z. B. um die Bedeutung der (Schaffung einer eventuell auch Starken) Künstlichen Intelligenz für den Menschen, ethische Aspekte im Umgang mit und bei der Kreierung von Künstlicher Intelligenz, Künstliche Intelligenz in der Wissenschaft, die Frage nach dem ontologischen Status (verschiedener Formen) der KI und die Beziehung zwischen künstlicher und menschlicher bzw. allgemein natürlicher Intelligenz geht.

2. Begriffsklärungen

Vor der Darstellung einzelner klassischer Topoi der Technikphilosophie und der Analyse ihrer Anwendung auf KI sind einige Begriffsklärungen geboten. Eine erste notwendige Differenzierung betrifft unmittelbar den Begriff der KI¹⁵. Hier sind (auf der obersten Ebene) vier unterschiedliche KI-Begriffe voneinander zu unterscheiden, die in zwei Begriffspaaren auftauchen – Schwache und Starke KI sowie Spezialisierte und Allgemeine KI. Bei der Schwachen KI kommt es lediglich zu einer

¹⁵ Einen guten Überblick über die verschiedene Verwendung von KI-Begriffen liefert z. B. Rosengrün (2021), Kap. 1.2, wobei es in der Literatur eine Vielzahl divergierender und teils sich widersprechender Begriffsverwendungen gibt. Populäre Beispiele von diesbezüglichen Begriffsklärungen im Zusammenhang mit KI finden sich z. B. in Bostrom (2017), S. 36–40, Lenzen (2019), S. 27–34, Russell/Norvig (2012), Kap. 26.1 sowie 26.2 und Kaplan (2017), Kap. 4, Tegmark (2017), S. 63–65. Die insbesondere philosophisch bedeutsame Debatte um Schwache und Starke KI wurde maßgeblich durch Searle (1980) mit dem Gedankenexperiment des „Chinesischen Zimmers“ befeuert. Besondere Schwierigkeiten hinsichtlich der Begriffsklärung(en) ergeben sich u. a. aufgrund der Tatsache, dass die im Kontext der KI-Forschung bedeutsamen Begriffe ‚Intelligenz‘, ‚Denken‘, ‚Kognition‘, ‚Geist‘, ‚Bewusstsein‘, etc. ihrerseits bis heute mit einem enormen Klärungsbedarf behaftet sind. Hier liegt es beispielsweise an der Philosophie des Geistes und den Kognitionswissenschaften, weitere Fortschritte zu erzielen.

Simulation des Denkens, ohne dass ein tatsächliches Denken stattfindet¹⁶ – dies ist bei Starker KI gegeben: Im Gegensatz zur Schwachen KI verfügt eine Starke KI über Geist bzw. Bewusstsein und ist der Lage, zu denken und zu verstehen. Eine Spezialisierte KI ist eine anwendungsbezogene KI, die ausschließlich zur spezifischen Problemlösung wie beispielsweise Bilderkennung, Sprachverarbeitung oder Routenplanung in der Lage ist. Wenngleich sich hier Anwendungsgebiete überschneiden oder sich auch neue Anwendungsgebiete bei einer bestehenden Spezialisierten KI aufgrund analoger Gebietseigenschaften ergeben können, ist die prinzipielle Beschränktheit auf einen konkreten Anwendungstyp für die Spezialisierte KI charakteristisch. Demgegenüber ist diese Bereichsspezifität bei der Allgemeinen KI nicht gegeben: Bei der Allgemeinen KI handelt es sich um eine generalisierte Form der Künstlichen Intelligenz, die z. B. Chataufgaben ebenso beherrscht wie das Schachspielen, Bilderkennung, Übersetzungen – und die generell über (annähernd) die gleichen beobachtbaren Fähigkeiten verfügt, die der menschliche kognitive Apparat zu leisten im Stande ist. Eine Sonderform der Allgemeinen Künstlichen Intelligenz ist das Konzept einer Superintelligenz, das einem insbesondere in trans- und posthumanistischen Kreisen begegnet¹⁷ und das maßgeblich von Irving Good (1916–2006) 1965 publizierter Vorstellung einer „Ultraintelligenten Maschine“, die die letzte Erfindung sei, die die Menschheit zu machen habe, geprägt ist¹⁸. Unter einer Superintelligenz ist dabei grundsätzlich eine Allgemeine Künstliche Intelligenz gemeint, deren kognitive Leistungen diejenigen der Menschen weit übertrifft. Mit Nick Bostrom, dem wohl führenden Vertreter in der gegenwärtigen Superintelligenz-Debatte, lassen sich drei (kombinierbare) Formen der Superintelligenz voneinander unterscheiden¹⁹: Eine schnelle Superintelligenz überbietet den Menschen in der Geschwindigkeit der ansonsten dem Menschen äquivalenten kognitiven Leistungen, eine kollektive Superintelligenz erreicht die Überlegenheit durch

¹⁶ Der Simulationsbegriff muss hierbei mit einer gewissen Vorsicht gebraucht werden. Die besonders prominent und wirkmächtig von Searle (1980) vertretene und im KI-Zusammenhang häufig vorzufindende Behauptung des generellen Unterschieds zwischen Simulation und Duplikation (des Denkens) erfuhr auch eine reichhaltige Kritik. Vgl. für einen diesbezüglichen Überblick z. B. Dresler (2009), Kap. 3.3.5.

¹⁷ Einen ersten Überblick hierzu liefern z. B. Loh (2018), Kap. 7.3, Mainzer (2019), Kap. 10.4 und Rosengrün (2021), Kap. 3.2.

¹⁸ Vgl. Good (1965).

¹⁹ Vgl. Bostrom (2017), Kap. 3.

das komplexe Zusammenwirken von einzelnen für sich genommen dem Menschen unterlegenen Systemen, und eine Qualitative Superintelligenz verfügt zusätzlich zu den vom Menschen bekannten Fähigkeiten über völlig neuartige Aspekte der Intelligenz.

Der Großteil der aktuellen KI-Forschung bezieht sich – entgegen dem ursprünglichen Ansinnen auf der berühmten, die moderne KI-Forschung unmittelbar einleitenden Dartmouth-Konferenz, auf die die Bezeichnung „Künstliche Intelligenz“ zurückgeht und bei der die Allgemeine KI im Fokus stand²⁰ – auf Spezialisierte KI, und alle uns derzeit zur Verfügung stehende KI ist eine spezialisierte und (unseres Wissens nach ...) schwach. Schwache und Spezialisierte KI einerseits sowie Starke und Allgemeine KI andererseits bezeichnen häufig denselben Gegenstandsbereich bzw. werden auch synonym verwendet, doch dies ist nicht zwingend. Wenngleich prima facie – in Analogie zum Menschen – (1) die Verbindung von Geist bzw. Bewusstsein und Allgemeiner KI naheliegt und die uns heute schon großflächig zur Verfügung stehende (2) Spezialisierte KI als schwach angesehen wird, ist prinzipiell auch (3) eine Spezialisierte KI mit Geist bzw. Bewusstsein oder (4) eine Allgemeine KI ohne Geist bzw. Bewusstsein denkbar. Insbesondere (1) und (4) sind angesichts der laufenden und zu erwartenden Entwicklungen ein philosophisch vieldiskutiertes Thema, das mitten in die Philosophie des Geistes und das sogenannte „schwierige Problem des Bewusstseins“²¹ – die Frage, warum wir Menschen über subjektive Erlebnisgehalte bzw. Qualia verfügen – führt²².

²⁰ Die interdisziplinär ausgerichtete Konferenz wurde von den KI-Pionieren John McCarthy (1927–2011), Marvin Minsky (1927–2016), Nathaniel Rochester (1919–2001) und Claude Shannon (1916–2001) organisiert und im Sommer 1956 mit weiteren Teilnehmern als zweimonatiger Workshop am Dartmouth College in Hanover (New Hampshire) mit dem Ziel, die Möglichkeiten der und Wege zu einer künstlichen Simulation der menschlichen Intelligenz zu eruieren, durchgeführt. Vgl. zur Dartmouth-Konferenz und ihrem Ansinnen sowie zur weiteren Entwicklung der KI-Forschung z. B. Kaplan (2017), S. 29–36, Lenzen (2019), S. 21–24, Mainzer (2019), Kap. 2.3 und Rosengrün (2021), Kap. 1.1.

²¹ Dieser Terminus geht auf Chalmers (1995) zurück.

²² Beispiele für die (kontroverse) philosophische Auseinandersetzung mit dem Thema KI und Bewusstsein sind u. a. Lenzen (2019), S. 132–134, Mainzer (2003), S. 120–124, Metzinger (2017), Kap. 7, Rosengrün (2021), Kap. 2.1 und Tegmark (2017), Kap. 8, gute Einführungen in die Philosophie des Geistes sind z. B. Beckermann (2008) und Brüntrup (2018).

Eine andere notwendige Klärung, die im Vorfeld weiterer Überlegungen vorgenommen werden muss, betrifft den Technik²³-Begriff²⁴. Der sprachliche Ausdruck „Technik“ begegnet uns heute im Alltag in unterschiedlichsten Zusammenhängen²⁵: Beispielsweise ist die Rede von Sozial- und Kulturtechniken, es werden Managementtechniken beworben, Mnemotechniken sollen dem Gedächtnis helfen, Elektrotechnik wird an den Hochschulen studiert, die Technik im Flugzeug muss gewartet werden und die Technik des Sportlers oder Musikers wird bewundert. Generell können unter den Begriff ‚Technik‘ ‚Verfahren/Prozesse als Schemata, Fähigkeiten zur Aktualisierung/Realisierung/,Umsetzung‘ dieser Verfahren, die Aktualisierung dieser Verfahren selbst und schließlich die Resultate dieser Aktualisierung²⁶ fallen, wobei im Rahmen einer ersten Differenzierung eine Unterscheidung zwischen engen, mittleren und weiten Technikbegriffen sinnvoll ist²⁷. Das entscheidende Kriterium der Zuordnung ist hierbei die Antwort auf die Frage, inwieweit (menschliche Handlungs-)Fähigkeiten zusätzlich zu den materiellen Endprodukten zur Technik zählen (sollen). Als Beispiel für einen mittleren Technikbegriff kann Günter Ropohls (1939–2017) Technikkonzeption herangezogen werden, nach der ‚Technik‘ „künstlich gemachte Gegenstände und menschliches Handeln umfasst, aber nur solches Handeln, das es mit Artefakten zu tun hat“²⁸.

‚Technik‘ kann aus einer systematischen Warte²⁹ als das „abgeklärte Ganze der Verfahren und Hilfsmittel des Handelns“³⁰ verstanden werden, wobei diese „Ver-

²³ Das Wort „Technik“ geht auf das altgriechische *techné* zurück, das – ursprünglich im Kontext des Handwerks – zur Bezeichnung einer auf das Hervorbringen abzielende Form des Könnens diente, wobei *techné* – wie das lateinische *ars* – auch die Kunst beinhaltete. Vgl. hierzu z. B. Gatzemeier (2004), Hubig (2006), S. 35f., Janich (2004), S. 214, Ropohl (2009), S. 38 und Zoglauer (2002a), S. 9.

²⁴ Vgl. zu den vielfältigen Zugängen zum Technikbegriff z. B. Heichele (2016), Kap. 2.1 und 2.2 sowie Heichele (2020a), S. 48–52.

²⁵ Vgl. z. B. Krohn (2002), S. 200 und Nordmann (2008), S. 11f.

²⁶ Hubig (2006), S. 229.

²⁷ Vgl. zu dieser Unterscheidung z. B. Heichele (2016), S. 19f., Hubig (2006), S. 28 und insbesondere Ropohl (2009), S. 29f.

²⁸ Ropohl (2009), S. 30.

²⁹ Vgl. zu den folgenden Ausführungen Heichele (2016), Kap. 2.2 und Heichele (2020a), S. 48–52.

³⁰ Gottl-Ottlilienfeld (1923), S. 8.

fahren und Hilfsmittel“ – entsprechend einem klassischen Ansatz in der Technikphilosophie seit der Antike³¹ – bewusst und planvoll eingesetzt werden müssen und keiner Willkürlichkeit unterliegen dürfen. So wird ‚Technik‘ zu einem vielschichtigen „Inbegriff der Mittel“³² und die engen, mittleren und weiten Technikbegriffe ergeben sich aus der Antwort auf die Frage, was als Mittel angesehen wird³³. Mittel erfordern die Bezugnahme auf Ziele bzw. Zwecke³⁴ – und Technik muss in einem weiten Sinne regelgeleitet sein, denn es ist in Abgrenzung von Zufallsprodukten oder Willkürhandlungen eine ihrer charakteristischen Eigenschaften, eine „Zweckrealisation wiederholbar und erwartbar [zu] machen“³⁵. In diesem Kontext ist ein (zumindest teilweises) Antizipieren des Verfüg- und Machbaren zentral und Technik ist – wie von der Antike bis zur Gegenwart betont – in prinzipieller Weise von einem Möglichkeitsraum abhängig³⁶: Die Identifikation der Disponibilität technischer Prozesse und Produkte ist eine notwendige Voraussetzung für Technik – und die Charakterisierung der Technik als „Inbegriff der Mittel“ erhält hinsichtlich des Mittelbegriffs eine dahingehende Präzisierung, dass neben aktuellen auch potentielle Mittel im Fokus liegen müssen.

Technik als (Inbegriff) der Mittel hat die zentrale Eigenschaft eines Mediums³⁷: Technische Formen der aktiven Welterschließung, -gestaltung und -deutung zeichnen sich dadurch aus, dass kein bezugsunabhängiger objektiver Gegenstand „Welt“ als solcher unmittelbar zugänglich ist, sondern dass die Welt stets vermittelt erschlossen wird³⁸. Technik schiebt sich gewissermaßen zwischen Subjekt und Objekt, wobei unter die „Objekte“ neben der materiellen Welt auch Ziele epistemischer und praktischer Natur fallen können: Um ein (Handlungs-)Ziel zu erreichen, wird vom (Handlungs-)Subjekt ein (technisches) Mittel verwendet. Eine besonders prä-

³¹ Vgl. z. B. Hubig (2006), Kap. 4.1.

³² Hubig (2006), S. 107.

³³ Vgl. generell zur Mittelhaftigkeit der Technik Hubig (2006), Kap. 4.1.

³⁴ Vgl. zum Zusammenhang von Technik und Zielen bzw. Zwecken z. B. Engel/Karafyllis (2004), S. 240f., Gallee (2003), S. 23–24, Gottl-Ottlilienfeld (1923), S. 8, Hubig (2006), Kap. 4.2 und Sachsse (1978), S. 9.

³⁵ Hubig (2006), S. 229.

³⁶ Vgl. hierzu sowohl aus systematischer als auch aus philosophiehistorischer Warte z. B. Heichele (2016), S. 22–39 und Hubig (2006), S. 28–53.

³⁷ Vgl. hierzu und im Folgenden insbesondere Heichele (2016), S. 27–33.

³⁸ Vgl. z. B. Cassirer (2007), S. 47–51, Hubig (2006), S. 15 und Mittelstraß (1996), S. 12–17.

gnante Darstellung der Technik als etwas Zwischengeschaltetes – in eine sehr ähnliche Richtung argumentiert Ernst Cassirer (1874–1945) bei seinen Überlegungen zur Technik mit Blick auf den Zusammenhang von Denken und Werkzeug³⁹ – findet sich bei Hans Sachsse (1906–1992):

Wir wollen als technisches Handeln ein Handeln bezeichnen, das einen Umweg wählt, weil das Ziel über diesen Umweg leichter zu erreichen ist. Technisches Vorgehen ist also nicht eine unmittelbare Aktion, sondern eine, die Mittel verwendet, die Mittel zwischenschaltet. Diese Mittel sind etwas anderes als das Ziel selber, und daher führen sie zunächst von dem Ziel fort, aber sie haben die Eigenschaft, daß durch ihre Vermittlung das Ziel leichter erreichbar wird⁴⁰.

Der für das technische Handeln charakteristische „Umweg“ ist also keine willkürliche Abkehr vom regulären Weg, sondern sie resultiert (für gewöhnlich) aus der Tatsache, dass das angestrebte Ziel nur auf diese Weise oder zumindest effizienter erreicht werden kann, wodurch der „Umweg“ zu einem „Umzu-Weg“ wird⁴¹. Technik ist somit ein Zwischenschalten von Mitteln, das mit Blick auf das Erreichen bestimmter Ziele der Ermöglichung, Sicherung und Erleichterung eines Handlungserfolgs dient. Das Bisherige zusammenfassend kann das Spezifische der Technik im Rahmen folgender pragmatischer Technikexplikation [TE]⁴² an drei Punkten festgemacht werden⁴³:

[TE-1] Technik ist eine Form des Handelns, bei der Handlungssubjekte bestimmte Mittel zu bestimmten Zielen und Zwecken in Relation setzen. Hierzu müssen mögliche Mittel als solche überhaupt erst erkannt werden.

³⁹ Vgl. Cassirer (2004), S. 158f.

⁴⁰ Sachsse (1978), S. 9.

⁴¹ Vgl. auch Gallee (2003), S. 26f.

⁴² Vgl. Heichele (2020a), S. 51.

⁴³ Diese drei Punkte entsprechen bis auf kleine Änderungen in der Formulierung im Wesentlichen den beiden ersten „Kriterien des Technikbegriffs“ Martin Arnold Gallees sowie dessen fünftem Kriterium; vgl. Gallee (2003), S. 30.

- [TE-2] Das Besondere des technischen Handelns liegt darin, dass das Ziel bzw. der Zweck nicht unmittelbar angegangen wird, sondern ein strategisch bedingter Umweg eingeschlagen wird, der letztlich das Erreichen des Ziels überhaupt erst ermöglicht oder zumindest effizienter erreichen lässt. Der Umweg wird so oft zum einzig möglichen „Umzu“-Weg.
- [TE-3] Der Erfolg technischen Handelns muss dahingehend transparent sein, dass die Gründe des Gelingens nachvollziehbar sind. Dadurch entstehen Wiederholbarkeit und Planbarkeit⁴⁴.

Dieser Technikbegriff ist in Übereinstimmung mit antiken (und immer noch vielen gegenwärtigen) Vorstellungen ein weiter, der Technik insbesondere als Mittel bzw. Medium versteht, wobei unter die (potentiellen) Mittel – das Gleiche gilt für Artefakte als Produkte der Handlungen – nicht nur materiale Dinge fallen, sondern auch intellektuale und soziale⁴⁵. Damit können schließlich drei Formen der Technik voneinander unterschieden werden: (1) Realtechnik, (2) Intellektualtechnik und (3) Sozialtechnik⁴⁶. Unter Realtechnik wird dabei – mit Friedrich von Gottl-Ottlilienfeld (1868–1958) – „ein Eingriff [...] in die sinnfällige Außenwelt, ob nun organischer oder anorganischer Natur“⁴⁷ verstanden, Intellektualtechnik ist „ein Eingriff [...] in eine intellektuelle Sachlage [...]“; so, daß z. B. alle Methodologie, aber auch die Technik [...] des Rechnens hierher gehört⁴⁸, und Sozialtechnik meint einen „Eingriff [...] in die Beziehungen zwischen den Handelnden; wie z. B. bei der

⁴⁴ Die rationale Nachvollziehbarkeit unterscheidet beispielsweise Technik von Magie und ihrer (vermeintlichen) Wirkweise. So sollen beispielsweise nach Arnold Gehlen (1904–1976) sowohl Technik als auch Magie dazu dienen, die Welt zum Vorteil des Menschen zu lenken. Vgl. z. B. Gehlen (1986), S. 96f. Gemeinsamkeiten und Unterschiede von Technik und Magie werden z. B. auch in Cassirer (2004), S. 152–164 behandelt.

⁴⁵ Vgl. hierzu und im Folgenden Heichele (2016), S. 31–33.

⁴⁶ Diese Dreiteilung ist lediglich pragmatischer Natur und erhebt keinen Anspruch auf Vollständigkeit. Hinsichtlich der Begrifflichkeiten geht diese Unterscheidung maßgeblich auf Gottl-Ottlilienfeld (1923), S. 9 zurück, der darüber hinaus noch die Ebene der Individualtechnik einführte. Vgl. zu dieser Dreiteilung z. B. auch Hubig (2006), S. 38f.

⁴⁷ Gottl-Ottlilienfeld (1923), S. 9.

⁴⁸ Ebd.

Technik des Kampfes, des Erwerbes, bei Rhetorik und Pädagogik, bei der Technik des Regierens und Verwaltens“⁴⁹.

Die Anmerkungen zum Naturbegriff sollen an dieser Stelle kurz ausfallen⁵⁰: Er wird im Folgenden vorwiegend als Reflexionsbegriff genutzt, wobei unter ‚Natur‘ – mit deutlichen Anleihen an die aristotelische Tradition⁵¹ – derjenige Bereich der Welt verstanden wird, der hinsichtlich seines Entstehens sowie mit Blick auf seine (konkret: naturgesetzliche) Daseins-, (Wechsel-)Wirkungs- und Entwicklungsform unabhängig vom reflektierenden Eingreifen⁵² ist bzw. sein kann⁵³. Die sich daraus ergebende Opposition zum Technikbegriff⁵⁴ ist dabei in erster Linie der reflexiven Funktion geschuldet: Die aristotelische Fundamentalunterscheidung zwischen Technik und Natur in einem ontologischen Sinne wird an dieser Stelle explizit zurückgewiesen⁵⁵: So bestimmt die Natur mit ihren Gesetzen nicht nur – wie (zumindest, sofern man von einer Fokussierung auf die Naturgesetze absieht) auch von Aristoteles erkannt – den Möglichkeitsraum für Technik, sondern Technik unterliegt z. B. selbstverständlich auch den Naturgesetzen und diese werden – zumindest vielfach – explizit genutzt.

⁴⁹ Ebd.

⁵⁰ Einen prägnanten Zugang zum Naturbegriff liefert Mittelstraß (2004).

⁵¹ Vgl. hierzu z. B. Heichele (2016), S. 22–24.

⁵² Damit ist zwar klassischerweise das (technische) Handeln des Menschen – bzw. generell eine Unabhängigkeit vom Menschen – gemeint, doch dies ist vor dem Hintergrund der bisherigen Ausführungen nicht zwingend: Prinzipiell ist auch ein technisches Handeln bei Tieren denkbar bzw. dieser Punkt wird kontrovers diskutiert. Vgl. zu dieser Problematik, die große Ähnlichkeit mit der erkenntnistheoretischen Debatte um Wissen bei (anderen hochentwickelten) Tieren aufweist, z. B. Heichele (2016), Fn. 114. Zudem würde eine vollständige Unabhängigkeit vom Menschen u. a. seine evolutionäre Natur leugnen.

⁵³ Eine vielschichtige Darstellung unterschiedlicher Naturbegriffe und Naturverhältnisse findet sich z. B. in Kirchhoff/Karafyllis (2017), eine Analyse des in unserem Zeitalter des Anthropozäns so bedeutsamen und oft sträflich vernachlässigten ökologischen Naturbegriffs liefert Soentgen (2020).

⁵⁴ Die in diesem Zusammenhang entstehenden Schwierigkeiten fundamentaler Begriffsexplikationen von ‚Natur‘ und ‚Technik‘ u. a. mit Blick auf biotechnologische Eingriffe des Menschen sowie z. B. hinsichtlich der generellen Tatsache, dass der Mensch und seine (technischen) Artefakte einen natürlichen Ursprung haben und damit in gewisser Weise ebenfalls Naturprodukte sind, werden hier nicht weiter thematisiert.

⁵⁵ Vgl. zur insb. in der (Früh-)Neuzeit feststellbaren Abkehr von der aristotelischen Annahme einer ontologischen Verschiedenheit von Natur und Technik Heichele (2016).

Mit diesen Vorarbeiten kann auch der Begriff der KI noch einmal präzisiert werden. Hierzu wird als Grundlage auf den Systembegriff rekurriert, wobei Systemen u. a. die Eigenschaften ‚künstlich‘ und ‚intelligent‘ zukommen können. Unter einem System ist dabei in erster Näherung schlicht eine geordnete, aus einzelnen in bestimmten Relationen zueinanderstehenden Elementen zusammengesetzte Ganzheit, die von ihrer Umwelt (trotz möglicher Wechselwirkungen) abgrenzbar ist und sich von dieser u. a. hinsichtlich der Komplexität unterscheidet, zu verstehen⁵⁶. ‚Künstlich‘ sind nun solche Systeme, deren Entstehung, Bestehen und Entwicklung sich einem technischen Eingreifen verdanken, und mit Blick auf ‚Intelligenz‘ gilt im Einklang mit Klaus Mainzers Arbeitsdefinition im Folgenden:

Ein System heißt intelligent, wenn es selbstständig und effizient Probleme lösen kann. Der Grad der Intelligenz hängt vom Grad der Selbstständigkeit, dem Grad der Komplexität des Problems und dem Grad der Effizienz des Problemlösungsverfahrens ab⁵⁷.

Mit ‚KI‘ ist also vor diesem Hintergrund – mit den bereits genannten Möglichkeiten der weiteren Differenzierung – in Abgrenzung zur natürlichen (z. B. menschlichen) Intelligenz ein technisch hergestelltes System, das eigenständig und effizient Probleme lösen kann, gemeint. Wenngleich diese Explikation keine Aussagen über den geplanten Weg derartiger Systemrealisierungen macht, ist in der Praxis – und das gilt auch für die folgenden Ausführungen – typischerweise der Informatik-lastige Ansatz einer (im weiten Sinne) Computer- bzw. Roboterintel-

⁵⁶ Ausführliche Darstellungen zum Systembegriff unter besonderer Berücksichtigung der Technik (Ropohl) und der Natur (Schurz) liefern z. B. Ropohl (2009), Kap. 3 und Schurz (2011), Kap. 7, eine Einführung in Komplexität und komplexe Systeme liefert Mainzer (2008). Sowohl die maßgeblich u. a. auf Ludwig von Bertalanffy (1901–1972) und Anatol Rapoport (1911–2007) zurückgehende interdisziplinäre (Allgemeine) Systemtheorie, die in den 1990ern zu einer Theorie der komplexen adaptiven Systeme wurde, als auch diverse Spezialisierungsformen sind nicht nur im Rahmen der Untersuchung natürlicher und technischer Phänomene ein bedeutendes (meta-theoretisches und meta-)wissenschaftliches Hilfsmittel, sondern spielen auch unmittelbar für die KI-Forschung eine wichtige Rolle. Vgl. zu komplexen Systemen aus Sicht der Wissenschaftstheorie z. B. Kuhlmann (2009).

⁵⁷ Mainzer (2019), S. 3.

lizenzen gemeint⁵⁸. Die schon auf einer ganz basalen Ebene (prinzipielle) potentielle multiple Realisierbarkeit der KI wird – mit Blick auf die öffentlichkeitswirksamen KI-Debatten – am ehesten im Zusammenhang mit Superintelligenz diskutiert⁵⁹.

3. KI und die Revolutionen im Mensch-Natur-Technik-Verhältnis

Die Menschheitsgeschichte ist geprägt von einer Vielzahl von Revolutionen, die das Mensch-Natur-Technik-Verhältnis betreffen⁶⁰: Eine genauere Analyse dieser Revolutionen macht dabei deutlich, wie sehr Neubetrachtungen und Innovationen der Technik – im eben explizierten weiten Sinn – unsere Stammes- und Kulturgeschichte beeinflusst(en)⁶¹. Vor ca. 40.000 Jahren begann mit der jungpaläolithischen Revolution die kulturelle Evolution: Durch rasante Fortschritte in (der Interaktion von) verschiedenen Bereichen der Real-, Intellektual- und Sozialtechnik – z. B. Werkzeuge und Waffen, Musikinstrumente und Kunstwerke – und ihre Tradierung kam es zu einer (partiellen) Abkopplung der kulturellen⁶² Evolution von der biologischen Evolution, aus der eine rapide Änderung der Lebensumstände resultierte, die vor ca. 12.000 Jahren in die neolithische Revolution mit Ackerbau und Viehzucht und dem Übergang vom Nomadentum zur Sesshaftwerdung mündete.

Vor ca. 2.500 Jahren kam es mit dem Beginn der Philosophie bzw. (Proto-)Wissenschaft zu einer intellektuellen Revolution, die u. a. das systematische Reflektieren des Spannungsverhältnisses von Mensch, Natur und Technik beinhaltete, wobei insbesondere unter dem maßgeblichen Einfluss von Aristoteles Technik tendenziell als etwas der Natur Entgegengesetztes und qualitativ Minderwertiges betrachtet

⁵⁸ Man beachte aber u. a. auch die im Kontext der starken Version des Paradigmas ‚Künstliches Leben‘ verfolgten Ansätze im Umfeld der Synthetischen Biologie. Einen guten Überblick zur Synthetischen Biologie liefern z. B. Pühler/Müller-Röber/Weitze (2011).

⁵⁹ Vgl. hierzu z. B. Rosengrün (2021), Kap. 3.2.

⁶⁰ Eine ausführlichere Darstellung dieser Revolutionen findet sich in Heichele (2020a), S. 52–58.

⁶¹ Einen ersten Zugang liefert z. B. Welsch (2007), umfangreiche Ausarbeitungen zur Kultur- und Stammesgeschichte des Menschen – mit unterschiedlichen Schwerpunkten – finden sich z. B. in Grupe (2012), Henke/Rothe (2003), Tomasello (2014) und Welsch (2012).

⁶² Mit ‚Kultur‘ ist hier schlicht das Tradieren von Techniken gemeint.

wurde⁶³. Eine weitere intellektuelle Revolution lässt sich in der Renaissance und mit dem Beginn der neuzeitlichen Philosophie sowie dem (damit zusammenhängenden) Aufkommen der neuzeitlichen (Natur-)Wissenschaft vor ca. 500 Jahren konstatieren: Weitergehende Reflexionen über die Beziehung von Mensch, Natur und Technik führten – u. a. festzustellen bei Cusanus (1401–1464), Leonardo da Vinci (1452–1519) und Galileo Galilei (1564–1642) – zu der Auffassung, dass Mensch, Natur und Technik (zumindest in vielerlei Hinsicht) denselben Gesetzen unterworfen sind. Nicht zuletzt daraus erwuchs ein enormer Technikoptimismus, wie er beispielsweise bei Francis Bacon (1561–1626) anzutreffen und für die Neuzeit über weite Strecken typisch ist: Der Mensch kann und soll Technik nutzen, um das Leben zu verbessern und Wissen zu generieren.

Mitte des 18. Jahrhunderts begann die erste von – bis heute – vier Industriellen Revolutionen, die jeweils vor dem Hintergrund technischer Neuerungen für große gesellschaftliche, wissenschaftliche und ökologische Umwälzungen sorgten⁶⁴. Die Erste Industrielle Revolution (ca. 1760–1840) ist u. a. mit der Spinnmaschine und dem mechanischen Webstuhl, sowie mit der Dampfmaschine und der Eisenbahn verbunden, die Zweite Industrielle Revolution (ca. 1870–1930) mit der Elektrotechnik und der Fließbandproduktion. Hier wurde der Übergang von der mechanischen Produktion zur Massenproduktion ermöglicht. Die Erste Digitale Revolution läutete (ab den 1960ern) die Dritte Industrielle Revolution ein, die u. a. von der zeitlichen Abfolge Großrechner, PC und Internet geprägt war – dies bedeutete beispielsweise eine Automatisierung der Produktion durch Informationstechnologie. In der Vierten Industriellen Revolution – beziehungsweise Zweiten Digitalen Revolution – (seit den späten 2010ern) kam es schließlich durch mobiles Internet, KI und eine generelle digitale Vernetzung u. a. zu einer selbstorganisierten Produktion, die auch unter dem Namen „Industrie 4.0“ bekannt ist.

Als Ergebnis dieser Revolutionen ist eine Welt entstanden, in der sowohl der Mensch als auch die gesamte Natur weitgehend technisch überformt sind. Von der öffentlichen Infrastruktur bis zum Haushalt, in der Arbeitszeit wie in der Freizeit –

⁶³ Vgl. hierzu und im Folgenden etwas ausführlicher Heichele (2020a), S. 53–56, sowie sehr ausführlich Heichele (2016).

⁶⁴ Diese Vierteilung ist in der Wissenschaft nicht sakrosankt. Vgl. zur Industriellen Revolution generell z. B. Cardwell (1997) und Ziegler (2013), sowie für ein engagiertes Plädoyer für die Vierteilung Schwab (2016).

der Mensch und seine unmittelbare Lebenswirklichkeit sind von den verschiedenen Formen der Technik geprägt. Dabei wird das Fortschreiten der Digitalen Revolution – insbesondere in Form von KI und (damit eng zusammenhängend) Big Data – in naher Zukunft zu längst angelaufenen fundamentalen Transformationen in u. a. Alltag, Wirtschaft, Politik, Wissenschaft und Kultur führen⁶⁵. Die technischen Handlungen des Menschen haben durch ihre Eingriffe in die Natur dabei in einem solchen Maße zu teilweise unumkehrbaren Veränderungen mit Blick auf biologische, geologische und atmosphärische Entwicklungen geführt, dass wir heute mit dem Anthropozän in einer neuen geochronologischen Epoche leben, die den Interventionen des Menschen geschuldet ist⁶⁶.

Bei einer genaueren Betrachtung wird deutlich, dass in der Ersten und Zweiten Industriellen Revolution vorwiegend eine Automatisierung der Realtechnik stattfand, in der Dritten und Vierten bzw. in der Digitalen Revolution zunehmend eine Automatisierung der Intellektual- und (gerade in der Vierten Industriellen bzw. Zweiten Digitalen Revolution, z. B. hinsichtlich diverser politischer oder ökonomischer Manipulationsverfahren im Kontext von Big Data) Sozialtechnik. Die gegenwärtig primär durchgeführten KI-Bestrebungen als Computer- bzw. Roboterintelligenz sind ein Streben nach einer maschinell ausgelagerten Automatisierung der Intelligenz, die ihrerseits eine (vormals nur natürlich vorkommende) Form der Intellektualtechnik ist.

4. Klassische technikphilosophische Topoi mit KI-Relevanz

Die Liste klassischer Topoi, die das philosophische Reflektieren der Technik (meist schon) seit den Anfängen in der Antike maßgeblich bestimmen und für

⁶⁵ Vgl. zur Digitalisierung generell z. B. Floridi (2015) und Wolff/Göbel (2018), zu KI Lenzen (2018) und Mainzer (2019) sowie zu Big Data Kolany-Raiser (2018) und Pietsch/Wernecke/Ott (2017).

⁶⁶ Vgl. zum Anthropozän einführend z. B. Leinfelder (2020) sowie zu einer Philosophie für das Anthropozän Heichele (2020b). Uwe Voigt (2020a) konstatiert für das Anthropozän (u. a.) eine „geistige Umweltkrise“ in Form eines (logischen) anti-universalistischen Narzissmus, der ein (zur Lösung der Probleme dringend benötigtes) vernunftgemäßes Reflektieren und Diskutieren erschwert bzw. gar verunmöglicht. Derartige Überlegungen können und müssen auch für das gerade erst beginnende Zeitalter der Digitalisierung angestellt werden.

KI relevant sind, ist lang⁶⁷. Die folgenden Anmerkungen befassen sich schwerpunktmäßig mit einer kleinen Auswahl derjenigen, die das Mensch-Natur-Technik-Verhältnis betreffen – müssen jedoch an dieser Stelle aus Platzgründen in einem sehr kursorischen Rahmen verbleiben. Einige der Topoi korrespondieren unmittelbar mit den drei großen Zielen [KIZ], die die derzeitige KI-Forschung auszeichnen⁶⁸.

- | | |
|---------|--|
| [KIZ-1] | Konkrete Problemlösung mit Hilfe Spezieller KI, |
| [KIZ-2] | Verbesserung des Verständnisses menschlicher Intelligenz
bzw. allgemein menschlicher Kognition, |
| [KIZ-3] | Schaffung einer Allgemeinen KI in Anlehnung an den Menschen als Universalisten. |

So steht beispielsweise für [KIZ-1] u. a. der Topos der Überwindung menschlicher Unzulänglichkeiten durch Technik und der der Naturbeherrschung durch Technik, für [KIZ-2] der einer wechselseitigen Erkenntnisfunktion von Mensch, Natur und Technik und für [KIZ-3] der der technischen Nachahmung der Natur bzw. des Menschen als zweite Natur Pate. Diese sowie weitere Topoi werden im Folgenden etwas ausführlicher beleuchtet.

Das vielleicht wirkmächtigste und prominenteste – und in diesem Aufsatz auch am ausführlichsten behandelte – Paradigma bei der Bestimmung des Mensch-Natur-Technik-Verhältnisses ist die von der Antike bis zur Gegenwart (in oftmals spezifischen Kontexten und nur auf bestimmte Hinsichten abzielende) vertretene Auffassung, wonach Technik eine vom Menschen durchgeführte Nachahmung bzw. Imitation ist. Ein solcher Ansatz findet sich – mit teilweise sehr unterschiedlichen Detailausarbeitungen – u. a. bei Aristoteles, Hugo v. St. Viktor (1097–1141), Cusanus, Leonardo da Vinci, Francis Bacon, Ernst Kapp

⁶⁷ Einen ersten Überblick liefert z. B. Heichele (2020a), S. 58–62, ein überaus reichhaltiger Fundus lässt sich aus Gallee (2003), Kap. 2 und Kap. 3 extrahieren.

⁶⁸ Vgl. zu diesen Zielen Lenzen (2019), S. 31–34.

und Arnold Gehlen⁶⁹. Eine wichtige systematische Differenzierung ergibt sich, wenn man analysiert, was genau Gegenstand der Nachahmung oder Imitation ist bzw. sein soll: Hier stößt man auf funktionelle und strukturelle Analogien⁷⁰. Funktionsanalogien liegen dann vor, wenn die Natur hinsichtlich einer konkreten Funktion bzw. eines bestimmten Ziels [TE-1] als Vorbild für die Technik dient, Strukturanalogien, wenn auch der zur Zielerreichung gegangene (Umzu-) Weg [TE-2] von der Natur inspiriert ist⁷¹. Die rationale Nachvollziehbarkeit der Technik bzw. die Legitimation ihres Einsatzes zur Zielerreichung [TE-3] gründet bei Funktionsanalogien (v. a. in der Frühphase der technischen Forschung) auf dem Verweis auf das durch die erfolgte (Natur-)Realisation erwiesenermaßen prinzipiell Mögliche, bei den Strukturanalogien auf der Referenz auf den (in einem nicht intentional zu verstehenden) erfolgreich gegangenen Weg der Natur zu dieser Zielerreichung. Ein klassisches Beispiel einer starken Nachahmung durch Technik, die nicht nur mit Funktions-, sondern auch mit Strukturanalogien arbeitet, ist die auf Leonardo da Vinci zurückgehende⁷² Bionik⁷³: Eine direkte Entsprechung in der KI-Forschung sind (sehr energieeffizi-

⁶⁹ Vgl. zu einer detaillierten Darstellung Heichele (2016).

⁷⁰ Vgl. zu einem ersten Zugang zum Unterschied von Funktions- und Strukturanalogien z. B. Zoglauer (2002a), S. 14–19.

⁷¹ Die Kopplung der Struktur- an die Funktionsanalogie – sprich der Umstand, dass in der Natur vorzufindende Funktionen über imitierte Strukturen erreicht werden sollen – ist zwar der häufigste Fall, aber keine Notwendigkeit. Ein KI-relevantes Beispiel ist das nach John Hopfield benannte Hopfield-Netz, das als Künstliches Neuronales Netz nicht wie sonst (siehe unten) biologische Aspekte imitiert, sondern auf Erkenntnisse der theoretischen Physik im Zusammenhang mit dem Ferromagnetismus in Festkörpern rekurriert. Vgl. hierzu z. B. Mainzer (2019), S. 110–114. Dies korreliert mit dem technikphilosophischen Topos der Offenheit der Technik: Technik kann oftmals auch für andere als die ursprünglichen Ziele eingesetzt werden. Vgl. dazu Galle (2003), z. B. S. 31 sowie S. 80–83, der von „Technik ‚auf Vorrat‘“ spricht. Ein klassisches und die Wissenschaft revolutionierendes Beispiel ist das Teleskop, das Galilei entgegen der angedachten (z. B. militärischen) Verwendungsweise als wissenschaftliches Instrument zur Untersuchung von Himmelsobjekten einsetzte. Vgl. hierzu z. B. Heichele (2016), S. 224–228. Ein Beispiel für die Offenheit der Technik im Hinblick auf KI ist die Verwendung von ursprünglich für seismologische Untersuchung entwickelten Verfahren für Predictive Policing. Vgl. hierzu z. B. Mohler (2011).

⁷² Vgl. z. B. Heichele (2020a), S. 59.

⁷³ Einen hervorragenden Einstieg in die Bionik bietet Nachtigall (2002).

ente) neuromorphe Computersysteme, die sich unmittelbar am (menschlichen) biologischen Gehirn orientieren⁷⁴.

Der Topos der Naturnachahmung bzw. -imitation ist für das Streben nach KI prägend: Auf einer sehr allgemeinen Ebene kann die von der Natur hervorgebrachte Intelligenz (grundsätzlich sowohl beim Menschen als auch bei nicht-menschlichen Tieren) – oder in einem weiteren Verständnis allgemein bestimmte natürliche kognitive Fähigkeiten – als Vorbild für die KI-Forschung angesehen werden, wobei hier eine generelle Funktionsanalogie vorliegt. Strukturanalogien tauchen dann in vielfältigen Varianten auf, wenn ein Blick auf die konkreten Versuche der Realisation von KI geworfen wird. Das wohl offensichtlichste Beispiel sind die u. a. im Machine Learning eingesetzten Künstlichen Neuronalen Netze (KNN)⁷⁵. Diese sind eine (abstrakte) Nachahmung natürlicher neuronaler Netze, wie sie beispielsweise im menschlichen Gehirn vorzufinden sind⁷⁶. Dabei spielt konkret u. a. die aus den Neurowissenschaften bekannte und nach ihrem Entdecker Donald O. Hebb (1904–1985) benannte Hebb'sche Lernregel eine wichtige Rolle, die die synaptische Plastizität des Gehirns beschreibt: Lernprozesse gehen mit einer Stärkung der synaptischen Übertragung zwischen in diesem Zusammenhang besonders relevanten Neuronen(verbänden) einher. Eine weitere Strukturanalogie wird deutlich, wenn man z. B. das Reinforcement Learning als eine

⁷⁴ Vgl. zu neuromorphen Computern z. B. Mainzer (2019), Kap. 10.1.

⁷⁵ Vgl. zu KNN z. B. Mainzer (2019), Kap. 7, Rosengrün (2021), Kap. 1.3 und Russell/Norvig (2012), Kap. 18.7. Einen Zugang zu den auf unterschiedliche Wege setzenden verschiedenen Strömungen der KI-Forschung liefern z. B. Görz/Schmid/Wachsmuth (2014) und Rosengrün (2021), Kap. 1.1. In erster Näherung lassen sich auf einer obersten Ebene zwei unterschiedliche Ansätze unterscheiden: auf der einen Seite die Top-down-orientierte symbolische KI, die das explizite Abarbeiten logischer Schlussregeln in den Fokus rückt, auf der anderen Seite die Bottom-up-orientierte konnektionistische KI, die auf die Simulation (der Verbindung) von Neuronen(gruppen) setzt. Das (sehr heterogene) Machine Learning, das „Wissen“ aus „Erfahrung“ zu generieren versucht, ist in beiden Richtungen ein wichtiges Forschungsgebiet, wobei in jüngster Zeit insbesondere im Kontext des eine Sonderform des Machine Learning darstellenden Deep Learning, bei dem KNN mit mehreren Zwischenschichten verwendet werden, große Durchbrüche erzielt wurden.

⁷⁶ Gleichzeitig muss jedoch betont werden, dass sich KNN trotz vieler Analogien in entscheidenden Punkten von natürlichen neuronalen Netzen unterscheiden. Ein Beispiel für ein Verfahren, das zwar von großer Relevanz für KI ist, aber – zumindest allem Anschein nach: vgl. demgegenüber jedoch z. B. Mainzer (2019), S. 115f. – keine biologische Entsprechung besitzt, ist Backpropagation: Dieser Lernmechanismus der schichtweisen Fehlerrückführung gilt gemeinhin als biologisch unplausibel.

der Methoden des Machine Learning betrachtet. Dieses Verstärkungslernen hat seinen Ursprung in der Psychologie bzw. Verhaltensbiologie und wird dort mit Blick auf Konditionierung behandelt.

Weitere Beispiele für die strukturelle Imitation sind u. a. bei der Zusammenführung von KI und Robotik auszumachen: Hier stehen beispielsweise Erkenntnisse aus der Evolutionsbiologie und Entwicklungspsychologie Pate und die aus der Philosophie des Geistes und der Kognitionswissenschaft bekannte These, wonach sich die menschliche Kognition (inkl. Bewusstsein) sowohl phylogenetisch als auch ontogenetisch maßgeblich vor dem Hintergrund der Wechselwirkung von Gehirn, Körper und (Um-)Welt entwickelt (hat) – „Developmental Robotics“, „Evolutionary Robotics“ und „Embodied Embedded Cognition“ sind an der Stelle nur einige wesentliche Schlagworte⁷⁷. Bei den – auch in diesem Zusammenhang durchgeführten – Überlegungen und Bestrebungen, eine Starke KI – also eine KI mit Geist bzw. Bewusstsein – zu schaffen, handelt es sich prinzipiell um eine Naturnachahmung im Sinne einer Funktionsanalogie, die allerdings vielfach – z. B. im Falle des Rückgriffs auf (im weiten Sinne) evolutionsbiologische Erkenntnisse – auch Strukturanalogien als zielführend beinhalten⁷⁸. Generell ist – wie bereits kurz angeklungen – die als „Leib-Seele-Problem“ bekannte Frage nach der Entstehung des Bewusstseins bzw. die nach dem Zusammenhang von Körper und Geist oder Physischem und Mentalem, die beispielsweise Qualia und Intentionalität betrifft, trotz aller Fortschritte in den (u. a.) Neurowissenschaften und der Philosophie des Geistes⁷⁹ nach wie vor ungeklärt und stellt die (neben der Entstehung des Universums) wohl größte epistemische

⁷⁷ Vgl. zur Zusammenführung von KI und Robotik z. B. Lenzen (2019), Kap. 3 und Mainzer (2019), Kap. 8, gute Hinführungen zu dieser Thematik aus Sicht der Philosophie und der Kognitionswissenschaft sind z. B. Fingerhut/Hufendiek (2017) und Lyre (2013).

⁷⁸ Vgl. zu der komplexen Problematik des (möglichen) Bewusstseins bei KI z. B. Lenzen (2019), S. 129–134, Rosengrün (2021), Kap. 2.3 und Tegmark (2017), Kap. 8, sowie zum Bewusstsein generell Metzinger (2005). Für eine allgemeine Annäherung an die Philosophie des Geistes sei noch einmal auf Beckermann (2008) und Brüntrup (2018) verwiesen.

⁷⁹ Eine Lösung des Leib-Seele-Problems kann nur aus der Philosophie (unter Bezugnahme auf die empirischen Wissenschaften) kommen: Empirische Wissenschaften gelangen (ohne Philosophie) nie über den Status der empirischen Adäquatheit ihrer Theorien hinaus – die theoretische Unterbestimmtheit der Empirie erlaubt keine Antwort auf derartige metaphysische Fragen. Vgl. dazu z. B. Heichele (2020c), S. 15f.

Herausforderung der Menschheit dar⁸⁰: Sämtliche gängigen Positionen in der Philosophie des Geistes – beispielsweise von Dualismus und Identitätstheorie über Funktionalismus bis hin zum nichtreduktiven Materialismus und Panpsychismus⁸¹ – haben mit (teilweise großen) Schwierigkeiten zu kämpfen, sind aber (mit mehr oder weniger umfangreichen Ad-hoc-Annahmen) gleichzeitig mit den Erkenntnissen der empirisch arbeitenden Neuro- und Kognitionswissenschaften kompatibel⁸² – und sind a priori offen für eine Starke KI. Argumente für oder gegen die prinzipielle Möglichkeit einer KI mit Bewusstsein werden im Kontext bestimmter Annahmen geführt und sind nicht automatisch an eine spezifische Position in der Philosophie des Geistes gekoppelt. Prima facie liegen u. a. zwei – sich auf den ersten Blick möglicherweise widersprechende – Grundüberlegungen nahe:

1. Eine (auch unter Berücksichtigung anderer Disziplinen die KI-Forschung letztlich maßgeblich bestimmende) informatische Zugangsweise kann den Phänomenen des Bewusstseins nicht gerecht werden, da hier bereits aus begrifflichen und systemgesetzlichen Gründen die subjektive Innenperspektive mentaler Zustände ausgeblendet wird.
2. Da die Natur eine „starke“ natürliche Intelligenz – sprich Bewusstsein – hervorgebracht hat, muss eine (technische) Wiederholung bei entsprechender Strukturgleichheit prinzipiell möglich sein.

Eine voneinander unabhängige (aber kombinierbare und an einigen Stellen auch verwandte) Auflösung dieses vermeintlichen Widerspruchs ergibt sich

⁸⁰ Es gibt auch Stimmen – siehe z. B. McGinn (1989) – die eine Lösung des Leib-Seele-Problems als prinzipiell für den Menschen kognitiv unzugänglich postulieren.

⁸¹ Der (zwar auch von einigen – wenigen – Philosophen vertretene, aber philosophisch dennoch völlig abwegige) eliminative Materialismus ist hier aus naheliegenden Gründen nicht genannt: Wer die Existenz eines natürlich entstandenen Bewusstseins leugnet, braucht sich nicht um ein künstlich hergestelltes Bewusstsein als Nachahmung Gedanken zu machen.

⁸² Vgl. zu dem häufig anzutreffenden Irrglauben, wonach ein – an dieser Stelle synonym gebraucht – Naturalismus/Materialismus/Physikalismus in besonderer Weise rational sei oder aus der (oder einem Bekenntnis zur) Wissenschaft folgen würde allgemein und einfüßend z. B. Tetens (2013), mit Blick auf die Philosophie des Geistes und logisch pointiert Meixner (2008).

u. a., wenn die Grundbausteine der Welt – und damit auch die der Systeme der Informatik – um eine qualitative Geist-Komponente als Innenseite sämtlicher natürlicher Entitäten erweitert werden⁸³, oder der Fokus auf emergentistische⁸⁴ Selbstorganisationsprozesse komplexer Systeme gerichtet wird⁸⁵, wobei hier ein Verstehen im strengen Sinn nicht gegeben (und je nach Konzeption auch gar nicht möglich) ist und ein alle Details kennender Konstrukteur einer Starken KI nicht benötigt würde: Nicht nur in diesem Fall⁸⁶ gerät die Klassifikation einer unter diesen Bedingungen entstandenen Starken KI als Technik ins Wanken⁸⁷, da hier der für Technik elementare Aspekt der rationalen Nachvollziehbarkeit [TE-3] nicht mehr gegeben ist⁸⁸.

Generell gibt es angesichts des bis heute ungelösten Leib-Seele-Problems und unserer damit zusammenhängenden Unkenntnis bezüglich der Entstehung des Bewusstseins im Zusammenhang mit einer Starken KI folgende Möglichkeiten [KIB]:

⁸³ Hier sei insbesondere auf den in jüngster Zeit in der Philosophie des Geistes florierenden Panpsychismus verwiesen – vgl. einführend z. B. Bischoff (2019) sowie zu einer umfangreichen Bestandsaufnahme Brüntrup/Jaskolla (2017). Eine – neben dem im Panpsychismus breit diskutierten Kombinationsproblem – besondere und in der Philosophie des Geistes allgemein oft vernachlässigte Herausforderung betrifft die Rolle und Entstehung der Vernunft. Vgl. zu dieser (auch im Kontext der Starken KI virulenten, aber hier nicht weiter behandelten) Problematik Voigt (2020b).

⁸⁴ Vgl. zur Emergenz (und ihren wissenschaftstheoretischen Problemen) einführend z. B. Hoyningen-Huene (2009), sowie zu einem ausführlichen Überblick z. B. Greve/Schnabel (2014).

⁸⁵ Wenngleich dieser Ansatz besondere Nähe zu klassischen Argumentationen im Rahmen des nichtreduktiven Materialismus aufweist, ist er nicht auf diese Position festgelegt.

⁸⁶ Dies betrifft grundsätzlich jede Form der Schaffung einer Starken KI, deren Entstehung epistemisch nicht vollständig nachvollzogen werden kann.

⁸⁷ U. a. in dieser Gemengelage tauchen z. B. metaphysische und ethische Fragen nach dem Personenstatus auf, die auch elementare juristische Implikationen besitzen.

⁸⁸ Der Umstand der Nicht-Nachvollziehbarkeit ist (mit den verschiedensten Konsequenzen) je nach Betrachtungsebene bereits bei heutigen KI-Systemen gegeben, bei denen die KNN eine „Black Box“ sind, deren genaue Prozesse sich unserer Kenntnis entziehen. Vgl. dazu z. B. Mainzer (2019), Kap. 11.1. Dem Umstand der Nicht-Nachvollziehbarkeit wird in der KI-Forschung u. a. durch das Streben nach sogenannter „Explainable Artificial Intelligence“ begegnet.

- [KIB-1] Wir leugnen die Möglichkeit der Entstehung einer Starken KI, diese ist aber möglich. Dabei kann es sein, dass wir mit unseren technischen Verfahren
- (a) dennoch keine Starke KI kreieren (können),
 - (b) eine Starke KI kreieren und dies nicht merken.
- [KIB-2] Wir sehen eine Starke KI als möglich an und sie ist es auch. Dabei kann es sein, dass wir mit unseren technischen Verfahren
- (a) tatsächlich eine Starke KI kreieren und diese auch als solche erkennen,
 - (b) eine Starke KI kreieren und dies nicht merken,
 - (c) keine Starke KI kreieren und dies auch so sehen,
 - (d) keine Starke KI kreieren, aber annehmen, wir hätten es getan.
- [KIB-3] Wir leugnen die Möglichkeit der Entstehung einer Starken KI und haben recht.
- [KIB-4] Wir sehen eine Starke KI als möglich an, aber sie ist es nicht. Dabei kann es sein, dass wir mit unseren technischen Verfahren
- (a) nichts kreieren, das wir als Starke KI ansehen,
 - (b) etwas kreieren, das wir fälschlicherweise als Starke KI ansehen.

Philosophisch (sowie z. B. auch gesellschaftlich) problematisch sind die Fälle, in denen wir falsch liegen: [KIB-1] (b), [KIB-2] (b) und (d) sowie [KIB-4] (b), wobei [KIB-1] (b) und [KIB-2] (b) potentiell sehr gefährlich werden können⁸⁹. Bis heute gibt es keine zufriedenstellende Antwort auf die Frage, welche (konkreten) Kriterien hinsichtlich der Zuschreibung von Bewusstsein bei einer (dann) Starken KI gelten sollen⁹⁰. Dieses Problemfeld ist eng zusammenhängend mit dem aus der Erkennt-

⁸⁹ Ein klassischer technikphilosophischer Topos betrifft die Eigendynamik sowie die (damit zusammenhängenden) nicht-intendierten Nebenfolgen technischer Entwicklungen. Vgl. dazu z. B. Gallee (2003), S. 48f.

⁹⁰ Vgl. hierzu z. B. Birnbacher (2005), Lenzen (2019), 129–134, Mainzer (2019), Kap. 7.3 sowie Rosengrün (2021), Kap. 2.1. Ein bis heute in der Literatur stark vernachlässigter Aspekt betrifft die Frage der Hypnotisierbarkeit – nicht als notwendige oder hinreichende

nisttheorie und der Philosophie des Geistes bekannten sogenannten „Problem des Fremdpsychischen“⁹¹: Letztlich hat ein jeder von uns nur auf sein eigenes (Selbst-) Bewusstsein direkten epistemischen Zugriff und wir wissen streng genommen nur von uns selbst, dass wir uns in bestimmten mentalen Zuständen befinden. Mit Blick auf die anderen (Menschen, aber auch viele nicht-menschliche Lebewesen) lässt sich unmittelbar nur sagen, dass sie vor dem Hintergrund einer Extrapolation der eigenen Erste-Person-Perspektive ein (auch und gerade sprachliches) Als-ob-Verhalten an den Tag legen: Sie verhalten sich so, als ob sie Schmerzen haben, glücklich sind, an den letzten Urlaub denken, usw. Daraus schließen wir, dass auch sie sich in den verschiedensten mentalen Zuständen befinden – gestützt von weiteren Analogieschlüssen, die sich insbesondere auf evolutionär entstandene und klassischerweise primär physiologische strukturelle Ähnlichkeiten zwischen uns und den anderen beziehen. Der Weg, vor diesem Hintergrund angesichts der (im weiten Sinne) systemtheoretisch begründbaren Plausibilität einer (relativen) Substratunabhängigkeit und multiplen Realisierbarkeit vieler mentaler Phänomene⁹² eventuell eines Tages⁹³ in vergleichbaren Fällen (als nicht notwendige, aber zumindest hin-

Voraussetzung der Bewusstseinszuschreibung, aber je nach Testaufbau als (möglicherweise starkes) Indiz.

⁹¹ Vgl. zum Problem des Fremdpsychischen allgemein z. B. Welding (2002), zum Problem des Fremdpsychischen im Zusammenhang mit KI z. B. Dresler (2009), Kap. 3.4.5 und Rosengrün (2021), Kap. 2.1 – ein diesbezüglicher Klassiker ist Putnam (1960).

⁹² Vgl. zu einem solchen systemtheoretischen Rahmen einführend z. B. Mainzer (2008) und Schurz (2011). Die These der Substratunabhängigkeit ist – mit oftmals ausbaufähiger Reflexionstiefe – u. a. im Kontext transhumanistischer Ideale weit verbreitet. Vgl. z. B. Bostrom (2017) und Tegmark (2017). Die (deutlich engere) Annahme der multiplen Realisierbarkeit findet sich generell in vielen klassischen Strömungen der Philosophie des Geistes.

⁹³ Einen schwerwiegenden ethischen Einwand gegen das menschliche Streben nach der Erschaffung einer Starken KI liefert – z. B. in Metzinger (2017), S. 278–286 – Thomas Metzinger, der auf die Leidensfähigkeit einer starken KI verweist und beim Übergang von unbewusster zu bewusster KI aufgrund technischer (Konstruktions-)Unzulänglichkeiten nicht zu rechtfertigendes Leid eines solchen technischen Systems erwartet. In diesem Zusammenhang bietet sich der Hinweis auf eine weitere ethische und u. a. im Zusammenhang mit den technikphilosophischen Topoi der nicht-intendierten Nebenfolgen sowie dem der Verselbständigung der Technik – vgl. dazu z. B. Heichele (2020a), S. 61 – stehende ethische Problematik an, die für gewöhnlich im Kontext von (nicht zwingend starken) Superintelligenzen zur Sprache kommt; vgl. z. B. Bostrom (2017) und Tegmark (2017) sowie für einen generellen Überblick Rosengrün (2021), Kap. 3.2: Superintelli-

reichende Bedingung) bei künstlichen Systemen Bewusstsein zu unterstellen und eine Starke KI zu konstatieren, ist nicht weit – und es bräuchte besondere Gründe, diese zu negieren.

Ein weiterer klassischer und mit der Imitationsthese eng verbundener technikphilosophischer Topos ist der einer – im Idealfall nicht nur einfachen, sondern doppelten bzw. – wechselseitigen Erkenntnisfunktion im Verhältnis von Mensch, Natur und Technik⁹⁴. Dieser ist u. a. bei Leonardo, Kapp und Cassirer anzutreffen. Hinsichtlich der KI ist jener Topos, dem die insbesondere in der Neuzeit gewachsene Erkenntnis einer basalen Strukturgleichheit der verschiedenen Gegenstandsbereiche zugrunde liegt, u. a. – siehe [KIZ-2] – im Kontext der wechselseitigen Befruchtung von Kognitionswissenschaft und KI-Forschung festzustellen⁹⁵. Die Entstehung beider läuft (in ihren modernen Ausprägungen) ab Mitte des 20. Jahrhunderts zeitgleich und mit starken Zusammenhängen ab, wobei insbesondere die Zuwendung zum (in der heutigen Philosophie des Geistes stark kritisierten) Funktionalismus, der nach den erkenntnistheoretischen und metaphysischen Beschneidungen im Behaviorismus und in der Identitätstheorie den Geist wieder wissenschaftlich salonfähig machte, eine wichtige Rolle spielt. Ebenso sind beide (Meta-)Disziplinen interdisziplinär ausgerichtet, wobei jede der jeweils anderen eine besondere Beachtung schenkt sowie u. a. (mit jeweils unterschiedlichen Gewichtungen) Erkenntnisse aus – in alphabetischer Reihenfolge – Anthropologie, Informatik, Linguistik, Neurowissenschaft, Philosophie und Psychologie berücksichtigt. Der gemeinsame Ausgangspunkt war das Ziel einer „Berechnung“ kognitiver Prozesse – in der Frühphase primär verstanden als symbolische Informationsverarbeitung. Einerseits zeigte sich dabei beim Versuch, Intelligenz bzw.

genzen könnten für die Menschheit (u. a.) dahingehend ein existenzielles Risiko bedeuten, dass sie unkontrollierbar werden und die Menschen unterwerfen. Unter der (gut begründbaren) Prämisse einer universell gültigen Vernunftethik – vgl. zu einem solchen Unterfangen unter gleichzeitiger Anerkennung evolutionswissenschaftlicher Erkenntnisse z. B. Illies (2006) – würden auch (zumindest) für eine hinreichend vernunftbegabte Superintelligenz leidminimierende und die Menschenwürde achtende Normen gelten.

⁹⁴ Vgl. hierzu z. B. Heichele (2016) im Allgemeinen, sowie dort speziell u. a. Kap. 2.3 sowie S. 111f. und ausführlich mit Blick auf Leonardo da Vinci Kap. 4 und 5.

⁹⁵ Vgl. zum Zusammenhang von Kognitionswissenschaft und KI-Forschung z. B. Lenzen (2019), Kap. 1, Russell/Norvig (2012), Kap. 1.1.2 und Sturm/Gundlach (2013), sowie zur Kognitionswissenschaft allgemein Stephan/Walter (2013) und zur Philosophie der Kognitionswissenschaften Walter (2017).

allgemein Kognition künstlich zu realisieren, die Notwendigkeit, bei der Erforschung menschlicher bzw. allgemein natürlicher kognitiver Phänomene begrifflich und logisch noch präziser zu argumentieren als bisher. Andererseits machte die immer genauere Analyse menschlicher Kognition deutlich, dass zu reduktionistische Ansätze der KI-Forschung der Vielschichtigkeit der menschlichen Kognition nicht gerecht werden können. Diese Einsichten führten in der Folgezeit zu ähnlichen Entwicklungen auf beiden Gebieten: Tendenziell ist eine Verschiebung von Top-down-Ansätzen einer starken computational-repräsentationalen Symbolverarbeitung hin zu einem Bottom-up-Ansatz einer dynamischen Selbstorganisation festzustellen.

Die notwendigen technikphilosophischen Reflexionen von KI unter Rückgriff auf bereits bestehende Problemstellungen – teilweise auch unter explizitem Bezug auf das Mensch-Natur-Technik-Verhältnis, gehen noch weit über das bis jetzt Gesagte hinaus: Weitere KI-relevante klassische Topoi⁹⁶ behandeln u. a. den Menschen als Mängelwesen, die menschliche Beherrschung der Natur durch Technik, die Entfremdung bzw. Loslösung von der Natur, die Gefahr, dass Technik nicht mehr als solche erkannt wird, die Abgabe von Verantwortung sowie drohende Autonomieverluste, die (sowohl gewollten als auch ungewollten) Erziehungsfunktionen der Technik, die Globalität und Irreversibilität (mancher) technischer Entwicklungen sowie den Symbolgehalt der Technik.

Literatur

- Beckermann, Ansgar (³2008): *Analytische Einführung in die Philosophie des Geistes*, Berlin–New York
- Berlinski, David (2000): *The Advent of the Algorithm: The 300-Year Journey From An Idea To The Computer*, San Diego – New York–London
- Birnbacher, Dieter (⁵2005): Künstliches Bewusstsein, in: Thomas Metzinger, Hg., *Bewußtsein: Beiträge aus der Gegenwartsphilosophie*, Paderborn, S. 713–729
- Bischoff, Alena (2019): Überall der Panpsychismus, in: Manfred Negele/Jan Levin Propach, Hg., *Geist-Erfahrung: Ein Beitrag zu einem Erfahrungsbegriff für die Geisteswissenschaften*, Würzburg, S. 43–60
- Bostrom, Nick (²2017): *Superintelligenz. Szenarien einer kommenden Revolution*, Berlin

⁹⁶ Vgl. z. B. Heichele (2020a), S. 58–62.

- Brüntrup, Godehard (2018): *Philosophie des Geistes: Eine Einführung in das Leib-Seele-Problem*, Stuttgart
- Brüntrup, Godehard/Jaskolla, Ludwig, Hg. (2017): *Panpsychism: Contemporary Perspectives*, Oxford
- Buchanan, Bruce G. (2005): A (Very) Brief History of Artificial Intelligence, in: *AI Magazine* 26/4, S. 53–60
- Cardwell, Donald (1997): *Viewegs Geschichte der Technik*, Braunschweig–Wiesbaden
- Cassirer, Ernst (2004): Form und Technik, in: ders., *Gesammelte Werke*. Herausgegeben von Birgit Recki. Band 17: Aufsätze und kleine Schriften (1927–1931), Hamburg, S. 139–183
- (2007): *Versuch über den Menschen. Einführung in eine Philosophie der Kultur*, Hamburg
- Chalmers, David (1995): Facing up the problem of consciousness, in: *Journal of Consciousness Studies* 2/3, S. 200–219
- Davis, Martin (2011): *The Universal Computer. The Road from Leibniz to Turing*, Boca Raton–London–New York
- Dresler, Martin (2009): *Künstliche Intelligenz, Bewusstsein und Sprache: Das Gedankenexperiment des »Chinesischen Zimmers«*, Würzburg
- Engel, Gisela/Karafyllis, Nicole C. (2004): Technik und Moderne, in: Gisela Engel/Nicole C. Karafyllis, Hg., *Technik in der frühen Neuzeit – Schrittmacher der europäischen Moderne*, Frankfurt am Main, S. 237–244
- Fingerhut, Joerg/Hufendiek, Rebekka (2017): Philosophie der Verkörperung: Ein Bericht über die Embodied-Cognition-Debatte, in: *Information Philosophie* 3, S. 16–32
- Floridi, Luciano (2015): *Die 4. Revolution. Wie die Infosphäre unser Leben verändert*, Berlin
- Gallee, Martin A. (2003): *Bausteine einer abduktiven Wissenschafts- und Technikphilosophie: Das Problem der zwei »Kulturen« aus methodologischer Perspektive*, Münster – Hamburg–London
- Gatzemeier, Matthias (2004): Techné, in: Jürgen Mittelstraß, Hg., *Enzyklopädie Philosophie und Wissenschaftstheorie*. Bd. 4 (Sonderausgabe), Stuttgart–Weimar, S. 214
- Gehlen, Arnold (1986): Die Technik in der Sichtweise der Anthropologie, in: ders., *Anthropologische und sozialpsychologische Untersuchungen*, Reinbek bei Hamburg, S. 93–103
- Good, Irving J. (1965): Speculations Concerning the First Ultrainelligent Machine, in: *Advances in Computers* 6, S. 31–88
- Görz, Günther/Schmid, Ute/Wachsmuth, Ipke (2014): Einleitung, in: Günther Görz/Josef Schneeberger/Ute Schmid, Hg., *Handbuch der Künstlichen Intelligenz*, München, S. 1–18
- Gottl-Ottlilienfeld, Friedrich v. (1923): *Grundriss der Sozialökonomik. II. Abteilung. Die natürlichen und technischen Bedingungen der Wirtschaft. II. Teil. Wirtschaft und Technik*, Tübingen
- Greve, Jens/Schnabel, Annette, Hg. (2014): *Emergenz: Zur Analyse und Erklärung komplexer Strukturen*, Berlin

- Grunwald, Armin/Julliard, Yannick (2005): Technik als Reflexionsbegriff: Überlegungen zur semantischen Struktur des Redens über Technik, in: *Philosophia naturalis* 42/1, S. 127–157
- Grupe, Gisela u. a., Hg. (2012): *Anthropologie. Ein einführendes Lehrbuch*, Berlin–Heidelberg
- Heichele, Thomas (2016): *Die erkenntnistheoretische Rolle der Technik bei Leonardo da Vinci und Galileo Galilei im ideengeschichtlichen Kontext*, Münster
- Heichele, Thomas (2020a): Das Spannungsfeld von Mensch, Technik und Natur: Von Ackerbau und Viehzucht zum Anthropozän, in: ders. (2020b), S. 47–66
- , Hg. (2020b): *Mensch – Natur – Technik: Philosophie für das Anthropozän*, Münster
- (2020c): Philosophie im 21. Jahrhundert, in: ders. (2020b), S. 9–24
- Henke, Winfried/Rothe, Hartmut (2003): *Menschwerdung*, Frankfurt am Main
- Hoyningen-Huene, Paul (2009): Reduktion und Emergenz, in: Andreas Bartels/Manfred Stöckler, Hg., *Wissenschaftstheorie. Ein Studienbuch*, Paderborn, S. 177–198
- Hubig, Christoph (2006): *Die Kunst des Möglichen I: Technikphilosophie als Reflexion der Medialität*, Bielefeld
- Illies, Christian (2006): *Philosophische Anthropologie im biologischen Zeitalter. Zur Konvergenz von Moral und Natur*, Frankfurt am Main
- Janich, Peter (2004): Technik, in: Jürgen Mittelstraß, Hg., *Enzyklopädie Philosophie und Wissenschaftstheorie*. Bd. 4 (Sonderausgabe), Stuttgart–Weimar, S. 214–217
- Kaplan, Jerry (2017): *Künstliche Intelligenz: Eine Einführung*, Frechen
- Kapp, Ernst (1978): *Grundlinien einer Philosophie der Technik: Zur Entstehungsgeschichte der Kultur aus neuen Gesichtspunkten. Neudruck mit einer Einleitung von Hans-Martin Sass*, Düsseldorf
- Kirchhoff, Thomas/Karafyllis, Nicole C. u. a., Hg. (2017): *Naturphilosophie. Ein Lehr- und Studienbuch*, Tübingen
- Kolany-Raiser, Barbara u. a., Hg. (2018): *Big Data und Gesellschaft. Eine multidisziplinäre Annäherung*, Wiesbaden
- Kornwachs, Klaus (2013): *Philosophie der Technik: Eine Einführung*, München
- Krohn, Wolfgang (2002): Technik als Lebensform. Von der aristotelischen Praxis zur Technisierung der Lebenswelt, in: Hans W. Ingensiep/Anne Eusterschulte, Hg.: *Philosophie der natürlichen Mitwelt: Grundlagen – Probleme – Perspektiven*, Würzburg, S. 193–210
- Kuhlmann, Meinhard (2009): Theorien komplexer Systeme: Nicht-fundamental und doch unverzichtbar?, in: Andreas Bartels/Manfred Stöckler, Hg.: *Wissenschaftstheorie. Ein Studienbuch*, Paderborn
- Leinfelder, Reinhold (2020): Das Anthropozän. Von der geowissenschaftlichen Analyse zur Zukunftsverantwortung, in: Thomas Heichele (2020b), S. 25–45
- Lenzen, Manuela (2019): *Künstliche Intelligenz: Was sie kann & was uns erwartet*, München
- Loh, Janina (2018): *Trans- und Posthumanismus zur Einführung*, Hamburg

- Lyre, Holger (2013): Verkörperlichung und situative Einbettung (embodied/embedded cognition), in: Achim Stephan/Sven Walter, Hg.: *Handbuch Kognitionswissenschaft*, Stuttgart
- Mainzer, Klaus (2003): *KI – Künstliche Intelligenz: Grundlagen intelligenter Systeme*, Darmstadt
- (2008): *Komplexität*, Paderborn
- (2019): *Künstliche Intelligenz – wann übernehmen die Maschinen?*, Berlin
- McGinn, Colin (1989): Can We Solve the Mind-Body Problem?, in: *Mind* 98/391, S. 349–366
- Meixner, Uwe (2008): Physikalismus, Dualismus und intellektuelle Redlichkeit, in: Markus F. Peschl/Alexander Batthyany, Hg.: *Geist als Ursache? Mentale Verursachung im interdisziplinären Diskurs*, Würzburg, S. 249–270
- Metzinger, Thomas, Hg. (2005): *Bewußtsein: Beiträge aus der Gegenwartsphilosophie*, Paderborn
- (2017): *Der Ego-Tunnel. Eine neue Philosophie des Selbst: Von der Hirnforschung zur Bewusstseinsethik*, München
- Mittelstraß, Jürgen (1996): *Leonardo-Welt. Über Wissenschaft, Forschung und Verantwortung*, Frankfurt am Main
- (2004): Natur, in: ders., Hg., *Enzyklopädie Philosophie und Wissenschaftstheorie*. Bd. 2 (Sonderausgabe), Stuttgart–Weimar, S. 961–964
- Mohler, George u. a. (2011): Self-Exciting Point Process Modeling of Crime, in: *Journal of the American Statistical Association* 106/406, S. 100–108
- Nachtigall, Werner (2002): *Bionik. Lernen von der Natur*, Berlin
- Nordmann, Alfred (2008): *Technikphilosophie zur Einführung*, Hamburg
- Pietsch, Wolfgang/Wernecke, Jörg/Ott, Maximilian, Hg. (2017): *Berechenbarkeit der Welt? Philosophie und Wissenschaft im Zeitalter von Big Data*, Wiesbaden
- Pühler, Alfred/Müller-Röber, Bernd/Weitze, Marc-Dennis, Hg. (2011): *Synthetische Biologie: Die Geburt einer neuen Technikwissenschaft*, Berlin–Heidelberg
- Putnam, Hilary (1960): Minds and machines, in: Sidney Hook, Hg., *Dimensions of mind, A symposium*, New York, S. 148–179
- Ropohl, Günter (2009): *Allgemeine Technologie: Eine Systemtheorie der Technik*, Karlsruhe
- (2010): Technikbegriffe zwischen Äquivokation und Reflexion, in: Gerhard Banse/Armin Grunwald, Hg.: *Technik und Kultur: Bedingungs- und Beeinflussungsverhältnisse*, Karlsruhe, S. 41–54
- Rosenberg, Jay F. (2009): *Philosophieren. Ein Handbuch für Anfänger*, Frankfurt am Main
- Rosengrün, Sebastian (2021): *Künstliche Intelligenz zur Einführung*, Hamburg
- Russell, Stuart/Norvig, Peter (2012): *Künstliche Intelligenz: Ein moderner Ansatz*, Hallbergmoos

- Sachsse, Hans (1978): *Anthropologie der Technik. Ein Beitrag zur Stellung des Menschen in der Welt*, Braunschweig
- Schurz, Gerhard (2011): *Evolution in Natur und Kultur: Eine Einführung in die verallgemeinerte Evolutionstheorie*, Heidelberg
- Schurz, Gerhard (*2014): *Einführung in die Wissenschaftstheorie*, Darmstadt
- Schwab, Klaus (2016): *Die Vierte Industrielle Revolution*, München
- Searle, John (1980): Minds, Brains and Programs, in: *Behavioral and Brain Sciences* 3/3, S. 417–457
- Soentgen, Jens (2020): Der ökologische Naturbegriff, in: Thomas Heichele (2020b), S. 115–130
- Stephan, Achim/Walter, Sven, Hg. (2013): *Handbuch Kognitionswissenschaft*, Stuttgart
- Sturm, Thomas/Gundlach, Horst (2013): Zur Geschichte und Geschichtsschreibung der ›kognitiven Revolution‹ – eine Reflexion, in: Stephan, Achim/ Walter, Sven (2013), S. 7–21.
- Tegmark, Max (2017): *Leben 3.0: Mensch sein im Zeitalter Künstlicher Intelligenz*, Berlin
- Tetens, Holm (2013): Der Naturalismus: Das metaphysische Vorurteil unserer Zeit?, in: *Information Philosophie* 3, S. 8–17
- Tomasello, Michael (2014): *Eine Naturgeschichte des menschlichen Denkens*, Berlin
- Turing, Alan (1950): Computing Machinery and Intelligence, in: *Mind* 49/236, S. 433–460
- Voigt, Uwe (2020a): Das Anthropozän als geistige Umweltkrise, in: Thomas Heichele (2020b), S. 85–102
- (2020b): Menschliche Personalität als Problem des Panpsychismus, in: *Salzburger Jahrbuch für Philosophie* 65, S. 145–158
- Walter, Sven (2017): Philosophie der Kognitionswissenschaft, in: Simon Lohse/Thomas Reydon, Hg.: *Grundriss Wissenschaftsphilosophie: Die Philosophien der Einzelwissenschaften*, Hamburg, S. 441–471
- Welding, Steen Olaf (2002): *Die Unerkennbarkeit des Geistes: Phänomenale Erfahrung und menschliche Erkenntnis*, Stuttgart
- Welsch, Wolfgang (2007): Just what is it that makes homo sapiens so different, so appealing?, in: *Deutsche Zeitschrift für Philosophie* 55/5, S. 751–760
- (2012): *Homo mundanus. Jenseits der anthropischen Denkform der Moderne*, Weilerswist
- Wolff, Dietmar/Göbel, Richard, Hg. (2018): *Digitalisierung: Segen oder Fluch? Wie die Digitalisierung unsere Lebens- und Arbeitswelt verändert*, Berlin
- Ziegler, Dieter (*2013): *Die Industrielle Revolution*, Darmstadt
- Zoglauer, Thomas (2002a): Einleitung, in: ders. (2002b), S. 9–45
- , Hg. (2002b): *Technikphilosophie. Texte*, Freiburg–München

Künstliche Intelligenz im Anthropozän? Aber natürlich!

Uwe Voigt

Zusammenfassung

Dieser Beitrag versucht, die Begriffe der Intelligenz, der Technik, der Künstlichen Intelligenz, des Anthropozäns und der Natur so zu klären, dass dies ihren wechselseitigen Zusammenhang erhellt. Intelligenz stellt sich dabei als die Fähigkeit dar, die Grenzen des eigenen Denkens als Probleme zu erkennen, was ein qualitatives und reflexives Problembewusstsein einschließt. Technik entspringt diesem Problembewusstsein und besteht in dem Versuch, Probleme zu lösen, um bestimmte gesetzte Zwecke zu erreichen. Künstliche Intelligenz ist demnach in Technik bereits inhärent, wird aber neuerdings als solche thematisiert. Dies gehört schon zum Problemkomplex des Anthropozäns als technischer Überformung der Umwelt auf diesem Planeten, was wiederum die Frage nach der Natur aufwirft, die hier relational verstanden wird als Wechselbezug möglicher subjektiver Standpunkte. Dies führt zu folgendem Ergebnis: „Natürlich“ kann Künstliche Intelligenz sein, insofern sie sich in einen derartigen Wechselbezug einfügen vermag, was im Anthropozän auch geraten zu sein scheint.

Abstract

This paper tries to clarify the concepts of intelligence, technology, artificial intelligence, the Anthropocene and nature so that their mutual interconnection is elucidated. In doing so, intelligence is understood as the ability to grasp the borders of one's own thinking as problems, which includes being qualitatively and reflectively aware of these problems. Technology stems from this awareness and consists in

the attempt to solve problems in order to reach certain set goals. That means that artificial intelligence is already inherent in technology, although it has been highlighted as such only recently. This already belongs to the complex problem of the Anthropocene, which is the technological transformation of the environment on this planet. That transformation, in turn, raises the question of nature, which here is taken to be relational, as an interconnection of possible subjective points of view. This leads to the following result: Artificial intelligence can be “natural” insofar it can integrate itself into such an interconnection, which also seems to be advisable in the Anthropocene.

1. Einleitung

Der Titel dieses Beitrags stellt eine beabsichtigte Provokation dar, denn er stellt zusammen, was allem Anschein nach nicht zusammengehört: Wie kann Künstliche Intelligenz natürlich sein? Und kann sie das überhaupt? Ist Künstliche Intelligenz als ein technisches Produkt nicht etwas, das gar nicht natürlich sein kann? Und ist aus eben diesem Grund die Rede vom Natürlichen gerade im Anthropozän nicht obsolet, da wir es in jener Epoche mit einer zutiefst von technischen Einflüssen geprägten Umwelt zu tun haben? Das „aber“ erhebt allerdings den Anspruch, dass all dies nichtsdestotrotz zusammengestellt, ja sogar zusammengedacht werden kann. Ist dieser Anspruch berechtigt, und wenn ja, wie kann er dies sein?

Jene Frage soll in diesem Beitrag durch eine Betrachtung der betreffenden Begriffe und ihrer wechselseitigen Zusammenhänge geklärt werden, was zugleich auch den weiteren Aufbau vorgibt. Dabei wird gar nicht erst versucht, jeweils *den* einschlägigen Begriff zu formulieren. Es wird vielmehr jeweils das Angebot *eines* solchen Begriffs gemacht, wenngleich verbunden mit der Hoffnung, dass sich damit erfassen lässt, was wir unter dem jeweils Gemeinten verstehen. Dieses Angebot kann hier schon allein aus Raumgründen nur skizzenhaft präsentiert werden.

Den Anfang macht dabei ein Begriff der Intelligenz, da er für das Verständnis einiger anderer Begriffe wichtig ist: für einen Begriff der Technik, die sich als eine bestimmte Anwendung von Intelligenz zeigen wird; damit auch für einen Begriff der Künstlichen Intelligenz, die wiederum auf einer derartigen Anwendung beruht, sowie für einen Begriff des Anthropozäns als technischer Überformung der Umwelt auf diesem Planeten. Die weitere Frage, ob und wie eine technisch über-

formte Umwelt im Allgemeinen und insbesondere Künstliche Intelligenz als ein Teil davon zugleich auch natürlich sein kann, führt dann zu einer Untersuchung des Begriffs des Natürlichen. Diese Untersuchung stellt sich der Herausforderung, dass der Begriff des Natürlichen mit dem hinter ihm stehenden Begriff der Natur notorisch problematisch zu sein scheint, und begegnet dieser Herausforderung, indem sie aus jenem Begriff einen reflexionslogischen Kern herauszuschälen versucht, wonach das Natürliche dasjenige ist, das wir benennen können, ohne es dafür beschreiben zu müssen. Am Schluss steht dann ein Plädoyer dafür, dass Künstliche Intelligenz im Anthropozän als etwas Natürliches begriffen werden kann und aus Gründen, die sich noch zu erweisen haben, auch als etwas Natürliches angestrebt werden sollte.

2. Ein Begriff der Intelligenz

Obwohl das Begriffswort „Intelligenz“ häufig inflationär und dabei ohne klar fasslichen begrifflichen Gehalt verwendet wird, scheint es dennoch erstaunlich einfach zu sein anzugeben, worin dieser begriffliche Gehalt besteht, wenn jenes Wort (oder eines seiner Synonyme) in terminologischer Weise angewendet wird: Dann ist unter Intelligenz nach einem verbreiteten Konsens die Fähigkeit zu verstehen, sich an Umweltmöglichkeiten auf bestmögliche Weise anzupassen¹. Aber eben dieser scheinbar einfache Zugang zeigt bei näherem Hinsehen, warum es in der Rede von Intelligenz so leicht zu begrifflicher Konfusion kommen kann: Wessen Fähigkeit und damit wessen Umwelt ist hier gemeint? Ist etwa bereits Wasser intelligent, wenn es sich Wege durch eine Landschaft bahnt und sich damit an die gegebenen Möglichkeiten dazu anpasst, diese Landschaft auf schnellstmögliche Weise zu verlassen? Gibt es eine Intelligenz der Gefühle („emotional intelligence“), die von derjenigen der Gedanken wesentlich verschieden ist? Und worin bestehen jeweils – oder gar allgemein – die Kriterien der Bewertung, die hier für die Weise der Anpassung vorzunehmen ist? Wäre es etwa allzu zynisch zu behaupten, das Aussterben einer Spezies sei immer noch die bestmögliche Weise der Anpassung an eine Umwelt, in der die Angehörigen jener Spezies nur ein leiderfülltes Leben hätten führen können? Jener Begriff der Intelligenz bedarf offenbar dringend einer Erläuterung.

¹ Vgl. Friedrich (2021), S. 146.

Eine solche, für den gegebenen Anlass weiterführende Erläuterung findet sich in der Wissenschaftslehre Johann Gottlieb Fichtes². Fichte denkt sich hier unter dem Titel des Ichs eine endliche Instanz, die Bewusstseinsinhalte hervorbringt und zugleich deren Träger ist. Ohne auf sich selbst zu reflektieren, würde das Ich einfach damit fortfahren, eben dies zu tun, Bewusstseinsinhalte hervorzubringen und sie für sich bewusst zu haben, und in diesem Sinne ein Handeln zu begehen, das in präreflexivem Denken besteht³, wenn es nicht auf einen auf seinem Weg liegenden Widerstand (auf transkribiertem Griechisch: ein *pro-blêma*) stieße, an der sich jene Tätigkeit bricht. Dieser Widerstand besteht darin, dass das Ich auf einen Bewusstseinsinhalt stößt, der ihm als etwas begegnet, das nicht von jenem Ich hervorgebracht worden ist. Etwas wird als etwas gedacht, was selbst nicht durch das Denken hervorgebracht worden ist. Das Denken stößt damit an eine Grenze – es denkt seine eigene Grenze – und wird dadurch auf sich selbst zurückgeworfen, wird reflexiv. Das Bewusstsein wird dadurch gleichsam in einem zu einem Selbstbewusstsein und einem Problembewusstsein: Ich bin auf ein *Problem* gestoßen; *ich* bin auf ein Problem gestoßen. Ein Problem ist dabei etwas, auf das das Ich sich beziehen kann und bei dem sich zugleich die Frage stellt, wie es sich darauf beziehen kann und wie es sich darauf beziehen soll. Ein dergestalt reflexiv und selbstreflexiv, zugleich aber auch intentional gewordenen⁴, auf eine von ihm unterschiedene Welt und damit im heutigen Sinne Umwelt bezogene Ich bezeichnet Fichte als Intelligenz⁵. Er verwendet diesen Ausdruck demnach als Bezeichnung des Trägers einer Eigenschaft, wie wir es heute auch und gerade im Fall der „Künstlichen Intelligenzen“ tun. Umso weniger stellt diese semantische Nuance ein Hindernis dafür dar, hier eine nähere Bestimmung auch des üblichen Verständnisses von Intelligenz als einer Eigenschaft zu erkennen: Intelligenz ist der erfolgreiche Umgang eines Ichs oder, dem heutigen Diskurs angemessen, eines Subjekts mit einem Problem im angegebenen Sinn, wobei sich die Maßstäbe des Erfolgs aus dem Denken eben jenes Subjekts selbst zu ergeben haben. Wie Uwe Meixner geltend macht⁶, gibt es Intelligenz in diesem Sinne nur bei einem und für ein Subjekt, nur als „Bewusstseinsintelligenz“, wobei unter

² Vgl. zum Folgenden Fichte (1794/2017).

³ Vgl. Borner (2016), Kap. 2.2.

⁴ Zu Fichte als einem der ersten Denker der Intentionalität vgl. Höhle (2018), S. 70.

⁵ Vgl. Fichte (1794/2017), S. 195f.

⁶ Vgl. Meixner (2021).

Bewusstsein der Umstand zu verstehen ist, dass es etwas gibt, das für dieses Subjekt als solches gegeben ist; dass es sich, wiederum modern gesprochen, für dieses Subjekt auf eine bestimmte Weise anfühlt, eben jenes Subjekt zu sein. Dementsprechend versteht Fichte unter elementaren Problemen in diesem Sinne einfache Erlebnisqualitäten (in neuerer Diktion: Qualia), an denen sich das Denken bricht, weil es sie so denken kann, dass es sie nicht hervorgebracht hat und auch nicht weiter zu analysieren vermag⁷. So verstandene Intelligenz hängt daher begrifflich mit der Fähigkeit zu qualitativem, ästhetischem Empfinden zusammen, was die Neigung dazu begründet, nichtmenschliche Wesen wie Androiden⁸ oder Bären⁹ ebenfalls als intelligent zu erachten, wenn ihnen diese Empfindungsfähigkeit zukommt.

Qualia als nicht weiter analysierbar zu denken bedeutet zugleich, sie als einfach zu denken, was Fichte mit dem Bild des geometrischen Punkts veranschaulicht. Damit lassen sich die von diesen Qualia ausgemachten Probleme auch quantifizieren, stellen sich als ein Problem unter vielen anderen dar. Mit Peirce lässt sich hier noch einen Schritt weitergehen¹⁰: Gerade aufgrund ihrer Einfachheit und damit ihrer Quantifizierbarkeit fungieren Punkte als Grenzen und damit als Verbindungen zwischen komplexen geometrischen Gebilden. Dementsprechend lassen sich auch komplexe Probleme beziehungsweise Problemkomplexe so auffassen, dass sie auf Zusammenhängen und Übergängen beruhen, die einen qualitativen Charakter haben, so dass Intelligenz im Umgang mit ihnen und damit auch zwischen ihnen, gleichsam „zwischen den Zeilen“¹¹, immer auch darauf beruht, dieses qualitativen Charakters inne werden zu können. Auch als problembezogenes Denken verstanden ist Intelligenz deshalb mit ästhetischem Empfinden verbunden. Dieses Empfinden verleiht den Problemen, um die es dabei jeweils geht, eine Relevanz, die sie gleichsam aus dem Hintergrundrauschen hervorhebt, und dem Subjekt, das mit ihnen konfrontiert ist, eine Motivation dazu, mit jenen Problemen umzugehen.

Der hier angebotene Intelligenzbegriff lässt sich so zusammenfassen: Intelligenz ist die Fähigkeit dazu, die Grenzen des eigenen Denkens als Probleme zu erkennen,

⁷ Vgl. Fichte (1794/2017), S. 230. Zur Anschlussfähigkeit dieser Überlegungen an die heutige Qualia-Debatte vgl. Henrich (2019), S. 156.

⁸ Vgl. St. Voigt (2021).

⁹ Vgl. Belke (2021).

¹⁰ Vgl. Pape (1997), S. 378–445; ders. (2006).

¹¹ Vgl. Dörner (2021), S. 218.

was auch ein qualitatives Bewusstsein von diesen Problemen einschließt, und sich auf dieser Grundlage sowohl auf jene Probleme als auch auf sich selbst als denkende Instanz zu beziehen.

3. Ein Begriff der Technik

Hier wird der von Thomas Heichele im Anschluss an Ernst Cassirer und Hans Sachsse erarbeitete Technikbegriff¹² herangezogen, und zwar nicht nur, weil er sehr klar dargelegt ist, sondern auch und gerade, weil er bestens zu dem gerade angedeuteten Verständnis von Intelligenz passt. Jenem Begriff zufolge lässt sich Technik nämlich primär als eine bestimmte Art und Weise intelligenten Handelns auffassen, genauer: als eine bestimmte Art und Weise, wie intelligentes Handeln mit sich selbst und seinen Problemen umgeht. Intelligentes Handeln richtet sich wie gesehen auf sich selbst und auf seine Probleme, wobei es durch seine Probleme auf sich selbst verwiesen wird, ist also reflexiv und intentional in einem. Der technische Umgang derartigen Handelns mit sich selbst und seinen Problemen besteht darin, diesen Bezug zu sich selbst und zu seinen Problemen als eine Zweck-Mittel-Relation zu begreifen, wobei der Zweck nicht im Handeln selbst und erst recht nicht in seinen Problemen gesehen wird, sondern in etwas, was gleichsam hinter den Problemen liegt. Denn die Probleme erscheinen hier als Hindernisse auf dem Weg zu einem Ziel, auf welches das Handeln gerichtet ist. Ist ein Problem mit Fichte veranschaulicht ein Punkt, an dem sich die fortlaufende Linie des präreflexiven Handelns bricht, dann ist das Ziel (der Zweck) ein weiterer Punkt, der hinter jenem ersten Punkt, dem Problem, liegt. Das intelligente Subjekt hat sich für sein Handeln eben jenes Ziel gesetzt, an das es aber aufgrund des Problems nicht gelangen kann, und sucht daher einen Weg, das Problem zu lösen. Wie Heichele ausführt, ist dieser Weg also zugleich ein Umweg und ein Um-zu-Weg. Das erste Mittel, zu dem das Handeln dabei greift, ist es selbst, beziehungsweise versteht es sich als ein Mittel, um weitere Mittel zu finden, die zum Erreichen des Zwecks führen. Der Zweck kann aber nur erreicht werden, wenn das Problem gelöst wird; solches Handeln stellt also

¹² Vgl. Heichele (2021), Abschnitt 2.

problemlösendes Denken par excellence dar¹³. Es richtet sich aus den genannten Gründen nicht unmittelbar auf seinen Zweck, sondern auf das zu lösende Problem. Die Mittel, die es dabei anwendet, um das Problem zu lösen und dadurch seinen Zweck zu erreichen, nennen wir für gewöhnlich ebenfalls Technik. Intelligentes Handeln als problemlösendes Denken vollzieht sich über derartige Mittel. Das intelligente Handeln setzt sich dabei selbst als ein Mittel ein, es nimmt also ebenfalls technischen Charakter an. In diesem Sinne bringt Technik stets auch künstliche Intelligenz mit sich: eine Intelligenz, die sich selbst technischen Charakter verleiht und dadurch einer bestimmten *technè*, einer Kunstfertigkeit, dient. Dies wird besonders in der sogenannten Intellektualtechnik deutlich, bei der das problemlösende Denken versucht, Probleme seiner eigenen Verfahrensweisen auf eine technische Weise (beispielsweise durch kontrollierte Anwendung einer formalisierten Logik) zu lösen.

Technik steht in einem ambivalenten Verhältnis zu den endlichen intelligenten Subjekten, die sie betreiben: Als endliche Subjekte können sie einerseits kaum umhin, zumindest einige Probleme technisch anzugehen und sich dadurch von ihnen zu entlasten¹⁴. Andererseits droht diese Entlastung gerade die zentralen Elemente derartiger Subjektivität zu untergraben: Je erfolgreicher Technik verfährt, desto mehr maskiert sie das Problem, zu dessen Lösung sie eingesetzt wird, und nimmt dadurch ihrem Subjekt auch den von diesem Problem gebotenen Anlass, sich seiner selbst reflexiv bewusst zu werden. Ist dieses Subjekt präreflexiv, solange es nicht durch Probleme zur Reflexion herausgefordert wird, kann es gleichsam in einen postreflexiven Zustand übergehen, wenn es mit Technik Probleme löst, ohne ihrer als solcher gewahr zu werden. Fiktive Szenarien, wie sie in der Populärkultur schon seit längerem und neuerdings markant in den Filmen der Terminator- und Matrix-Reihen entfaltet werden, lassen sich vor diesem Hintergrund als populärkultureller Reflex und damit als eine im Medialen steckengebliebene Reflexion auf die drohende Auslöschung oder zumindest Unterjochung endlicher (in diesen Fällen v. a. menschlicher) Subjektivität durch die von ihr eingesetzte Technik betrachten¹⁵.

¹³ Zu technischem Handeln als Problemlösen vgl. Popitz (1997), S. 106. Zur Technikgeschichte als einer Geschichte aufeinanderfolgender Problemlösungen und sich daraus wiederum ergebender Problemstellungen vgl. ders. (1995).

¹⁴ Zur Konzeption der Entlastung vgl. Claessens (2019).

¹⁵ Vgl. Irrgang (2020), S. 9–34.

Diese Drohung verschärft sich durch eine bestimmte Erscheinungsform von Technik¹⁶: Klassische Technik passt sich dem Problem an, zu dessen Lösung sie eingesetzt wird, und gestaltet sich dadurch jeweils auf eine spezifische, dem gesetzten Zweck und dem jeweiligen Problem angemessene Weise; sie ist, im Sinne dieser Angemessenheit und dadurch zumindest erstrebten Entsprechung, analog. Dadurch zerfällt klassische Technik zugleich in eine Mannigfaltigkeit unterschiedlicher Techniken, je nach den jeweils gesetzten Zwecken und den sich ihnen entgegenstellenden Problemen. Die Fülle der Techniken kann nun ihrerseits zu einem Problem werden, für das es wiederum eine technische Lösung zu geben scheint: die Entwicklung einer einzigen, einheitlichen Technik. Diese Technik passt sich nicht dem gegebenen Zweck und dem begnenden Problem an; vielmehr passt sie diese umgekehrt an sich an, macht sie möglichst unabhängig vom jeweiligen Bereich zu Gegenständen, die sie verarbeiten kann. Die Möglichkeit einer derartigen Technik gründet bereits in der Quantifizierung, die das Problembewusstsein vornimmt: Bei all ihrem je unterschiedlichen qualitativen Charakter begreift dieses Bewusstsein seine Probleme jeweils als verschiedene Einheiten. Der Ansatz der einen, einheitlichen Technik kann von daher darin bestehen, jene Einheiten als bloße Quantitäten zu verarbeiten, die an Fingern (Lateinisch: *digiti*) abgezählt werden könnten. Diese Technik lässt sich daher in einem weiten Sinn als digital bezeichnen. An der Schwelle zum 20. Jahrhundert hat sie einen großen Aufschwung nicht zuletzt dank der intellektualtechnischen Fortschritte auf dem Gebiet der Logik erfahren, auf dem es damals gelang, Quantität auf streng formalem Weg darzustellen¹⁷.

Gerade indem sie selbst zum Problem wird, kann Technik jedoch auch zur Selbstreflexion beitragen. Diese Selbstreflexion kann sich von den Polen des Subjekts und des durch Technik zu lösenden Problems her ergeben: vom Subjekt her, indem es sich seine gerade aufgezeigte ambivalente Beziehung zur Technik bewusstmacht, und vom Problem her, indem folgendem Zusammenhang Aufmerksamkeit zuteilwird: Technik richtet sich nicht unmittelbar auf den zu erreichenden Zweck, sondern auf ein Problem, das die Erreichung dieses Zwecks verhindert. Dadurch wird für die Technik die Bearbeitung jenes Problems selbst zu einem Zweck. Wie auch

¹⁶ Die hier vorgelegte Unterscheidung von analoger und digitaler Technik ist angeregt von der Differenzierung zwischen klassischer und moderner Technik bei Heidegger (1953/2000).

¹⁷ Vgl. Mainzer (2005); Davis (2011).

immer das Problem bearbeitet wird, richtet sich Technik als eine bestimmte Form intelligenten Handelns dabei immer auch auf dasjenige, womit das Problem zusammenhängt, und verändert auch diese Zusammenhänge, indem es das Problem bearbeitet¹⁸. Technik wird daher immer von Nebenwirkungen¹⁹ begleitet, die in der Verfolgung des jeweiligen Zwecks und der Bearbeitung eines damit verbundenen Problems nicht intendiert gewesen sind. Je mächtiger die jeweils eingesetzte Technik ist, desto gravierender können jene Nebenwirkungen werden. Auch die drohende Verselbstständigung von Technik gegenüber dem Subjekt, das sie betrieben hat, lässt sich als eine derartige Nebenwirkung verstehen, bei der sich der Zusammenhang zwischen Problem und Technik gegenüber demjenigen zwischen Technik und Subjekt als mächtiger erweist. Die Nebenwirkungen der Technik, die uns in Umweltfragen begegnen, tragen jedenfalls zur kritischen Reflexion über Technik in unserer Zeit bei.

4. Ein Begriff der Künstlichen Intelligenz

Dass Technik zumindest implizit immer auch schon in gewisser Weise Künstliche Intelligenz bedeutet, haben wir schon gesehen. Mitte des 20. Jahrhunderts explizierte sich dieser Zusammenhang, und zwar zunächst auf eine beiläufig wirkende Weise, indem der Ausdruck „artificial intelligence“ als ein „Schlagwort im Drittmittelantrag“ für eine einschlägige Konferenz geprägt wurde²⁰. Gemeint ist mit diesem Ausdruck und seinem deutschen Synonym „Künstliche Intelligenz“ jedenfalls ein sich technischer Einwirkung verdankendes Produkt, dessen Aktivitäten intelligentem Handeln entsprechen. Hierbei ergeben sich folgende Möglichkeiten: Die Aktivitäten jenes Produkts entsprechen entweder deshalb intelligentem Han-

¹⁸ Dies ist eine Spezialisierung der Einsicht Peter Sloterdijks, wonach jede Veränderung jeweils mehr verändert, als mit ihr eventuell beabsichtigt war, und dadurch unter Umständen lawinenartige Prozesse auslösen kann; vgl. Sloterdijk (1989), S. 23, 29. Auf dieser Grundlage kam Sloterdijk zu der Auffassung, dass das vom Menschen veränderte Erdsystem nunmehr nicht länger als neutrale Bühne verstanden werden kann, sondern als aktiver Mitspieler der Geschichte auftritt; vgl. ebd., S. 309. Damit nahm er die einschlägige These Bruno Latours (2017) um fast 30 Jahre vorweg.

¹⁹ Zum Begriff der Nebenwirkung vgl. Soentgen (2019), S. 45–49.

²⁰ Vgl. Rosengrün (2021b), Kap. 1.1.

deln, weil sie selbst Aktivitäten einer Intelligenz im oben angegebenen Sinne sind, also eines problembewussten Subjekts; dann ist von einer starken Künstlichen Intelligenz die Rede. Oder diese Aktivitäten entsprechen zwar intelligentem Handeln, sind aber selbst keine Aktivitäten einer derartigen Intelligenz; das fragliche Produkt agiert nur so, als ob es selbst intelligent wäre, ohne es aber zu sein. Dies ist das Charakteristikum einer schwachen Künstlichen Intelligenz. Ferner ist es denkbar, dass Künstliche Intelligenz Probleme beliebiger Art lösen kann und damit gleichsam die Vollendung der digitalen Technik im oben angegebenen Sinn bedeutet. In einem solchen Fall hätten wir es mit einer allgemeinen Künstlichen Intelligenz zu tun²¹. Alternativ dazu kann Künstliche Intelligenz ausschließlich dazu in der Lage sein, nur Probleme bestimmter Art zu lösen. Dies ist dann eine spezialisierte Künstliche Intelligenz, gleichsam in der Tradition der analogen Technik, auch wenn jene Künstliche Intelligenz mit „digitalen“ Mitteln herbeigeführt worden sein sollte. Diese Art Künstlicher Intelligenz ist es, die heute vielfach zum Einsatz kommt. Die Fragen, ob und wie starke Künstliche Intelligenz sowie allgemeine Künstliche Intelligenz möglich ist, bleiben notorisch offen. Nach den hier angestellten Überlegungen und angesichts der von Dietrich Dörner²², Sean McGrath²³ und Uwe Meixner²⁴ vorgebrachten Argumente darf allerdings zum einen angenommen werden, dass starke Künstliche Intelligenz nur auf einem Bewusstsein beruhen kann, das Qualitäten zu erfahren vermag und sich dadurch als Problembewusstsein zu gestalten vermag. Zum anderen legt dies nahe, dass eine allgemeine Künstliche Intelligenz zugleich auch eine starke Künstliche Intelligenz sein müsste, da das allgemeine Erkennen und entsprechende Bearbeiten von Problemen beliebiger Art offenbar auf einem Bewusstsein für diese Probleme und die von ihnen gebildeten Komplexe zu gründen hat. Eine mit phänomenalem Bewusstsein verbundene, gleichsam um es herum zentrierte starke Künstliche Intelligenz könnte auch von den von Sebastian Rosengrün zu Recht erhobenen metaphysischen Problemen verschont bleiben²⁵, die mit einer abstrakten und zugleich bewusstseinsfernen Konzeption Künstlicher Intelligenz verbunden sind.

²¹ Vgl. McGrath (2021), S. 56.

²² Vgl. Dörner (2021).

²³ Vgl. McGrath (2021).

²⁴ Vgl. Meixner (2021).

²⁵ Vgl. Rosengrün (2021a).

5. Ein Begriff des Anthropozäns

Ein vielbeachteter Ausdruck der Reflexion auf Technik, wozu der Begriff der Künstlichen Intelligenz verstärkten Anlass bietet, ist die Benennung der erdgeschichtlichen Situation, in der wir uns befinden, als Anthropozän²⁶. Zwar scheint hier zunächst nur von einem „neuen Zeitalter des Menschen“ die Rede zu sein. Doch äußert sich jenes „neue Zeitalter“ in den Auswirkungen, die das technische Handeln des Menschen in dessen Umwelt hinterlässt. Diese Auswirkungen sind nicht nur empirisch gut belegt; sie übertreffen in vielen Fällen sogar alle anderen Einflussfaktoren. Die technischen Mittel, die dies zustande bringen, manifestieren sich dabei selbst in jener Umwelt, ja verschmelzen geradezu mit ihr, wie dies Jens Soentgen exemplarisch in seiner Studie gezeigt hat, der zufolge sich der für viele Gewässer mit ähnlichem Schicksal stehende Fluss Lech in einen Cyborg, ein Mischwesen mit technischer Infrastruktur, verwandelt hat²⁷. Als ein derartiges Mischwesen entfaltet der Lech, wie Soentgen zeigt, allerdings auch Aktivitäten, die bei den menschlichen Eingriffen, die ihn geprägt haben, gar nicht beabsichtigt waren; und auch dabei steht er als ein Teil für ein Ganzes, eine planetare Umwelt, die sich aufgrund der Auswirkungen menschlicher Technik allem Anschein nach zunehmend destabilisiert, wofür die zahlreichen „jockey sticks“ sprechen, exponentiell anwachsende Messwertkurven, in denen sich jene Auswirkungen widerspiegeln. Technik wird dabei so sehr zum prägenden Faktor, dass ein eher räumlich orientierter Parallelbegriff zum Zeitraum des Anthropozäns trefflich als derjenige der Technosphäre geprägt worden ist²⁸: das komplexe System, das von der Technik gebildet wird, und das sich zunehmend auch ohne Rücksicht auf menschliche Interessen selbst organisiert, weil es insbesondere auf den Nebenfolgen der Technik beruht. Wenn sich die davon betroffene Biosphäre selbst als ein Erdsystem auffassen lässt, das organischen Charakter hat und von dem Naturwissenschaftler James Lovelock als „Gaia“ bezeichnet wurde²⁹, dann erleben wir gerade, wie Gaia von der Technosphäre durchdrungen und assimiliert wird – wie sich auch hier, im plane-

²⁶ Vgl. dazu Renn/Scherer (2017); Crutzen (2019); Horn/Bergthaller (2019); Heichele (2020).

²⁷ Vgl. Soentgen (2014).

²⁸ Vgl. Klingan/Rosol (2019).

²⁹ Vgl. Lovelock (2016).

taren Maßstab, ein Cyborg formiert, der zunehmend Eigenaktivitäten entfaltet³⁰. Diese Eigenaktivitäten richten sich nach den Zwecken jener Entität und haben es dabei mit entsprechenden Problemen zu tun, wobei nicht gewährleistet ist, dass diese Zwecke auch unsere Zwecke sind und dass zumindest einige unserer Zwecke nicht auch zu jenen Problemen zählen könnten.

Das Anthropozän lässt sich von daher als die technische Überformung der Umwelt auf diesem Planeten verstehen, und zwar in einem dreifachen Sinn: Es ist eine *technische* Überformung, ein von Technik hervorgerufenen Vorgang; es ist zudem eine *technische Überformung*: das hier Geformte nimmt selbst die Form von Technik an; und es ist schließlich auch eine *technische Überformung*, die sich an Zwecken und Problemen orientiert, die nicht mehr diejenigen derer sind, es zumindest nicht sein müssen, deren technischer Eingriff zu jener Situation geführt hat. Was diese Situation weiter kompliziert, ist der Umstand, dass sie, wie jeder Problemzusammenhang, von einer bestimmten qualitativen Beschaffenheit ist, es sich also auf eine bestimmte Weise anfühlt, in ihr zu sein. Angesichts vielfältiger Phänomene auf unterschiedlichen Ebenen von der individuellen Erfahrung bis hin zu internationalen Entwicklungen lassen den Verdacht zu, dass es sich dabei um eine Atmosphäre des logischen Narzissmus handelt – der Identifikation der betroffenen Subjekte mit dem Standpunkt, den sie jeweils einnehmen, und dem damit verbundenen zwanghaften Festhalten an jenem Standpunkt³¹. In einer derartigen Atmosphäre kann sich eine „Ökologie der Angst“ verbreiten³², die menschliche wie nichtmenschliche Subjekte in eine unfreiwillige Gemeinschaft der Leidenden zusammenzwingt³³.

Aufgrund der Maßstäbe, die diese Situation einnimmt, lässt sie sich von menschlicher Warte aus zwar einerseits umreißen, wie es hier gerade geschieht, und andererseits in einer Fülle von Messungen erfassen, worum sich empirische Wissenschaften bemühen. Für die Verquickung zwischen empirischen Aspekten und qualitativen Gehalten in jener Situation spricht zudem, dass die Lage im Anthropozän zunehmend auch zum Untersuchungsgegenstand von Geistes- und

³⁰ Vgl. Verf. (2020a); ders. (2020b).

³¹ Vgl. ebd.

³² Vgl. Soentgen (2018).

³³ Vgl. Dufresne (2019).

Kulturwissenschaften wird, den Environmental Humanities³⁴, die sich dem kulturellen Reflex von Umweltbedingungen widmen³⁵. In einer solchen auf vielfache Weise komplexen Situation benötigen wir offenbar jede Art von Intelligenz, die uns beim Erfassen und Bewältigen der entsprechenden Problemlagen helfen kann, und daher ebenfalls Künstliche Intelligenz, die dank ihrer überragenden Fähigkeit zur Datenverarbeitung sogar das Mittel der Wahl im Anthropozän zu sein scheint³⁶. Darüber hinaus könnte sich Künstliche Intelligenz in ihrer Verbindung mit weiteren technischen Mitteln als ein mächtiger Akteur erweisen, der uns angesichts der mit dem Anthropozän verbundenen Problemen zur Hilfe kommt, ja jene Situation eventuell sogar zugunsten eines neuen, von freundlichen Cyborgs gestalteten Erdzeitalters zu einem glücklichen Ende gelangen lässt³⁷. Doch gibt es nach den gerade angestellten Überlegungen Grund dazu, an diesem zunächst tröstlich wirkenden Szenario von James Lovelock zu zweifeln: Wie jedes andere technische Produkt, so ist auch Künstliche Intelligenz selbst Teil der Technosphäre und damit Teil des Problemzusammenhangs, mit dem wir es hier zu tun haben. Auch wenn das technisch überformte Erdsystem sich nicht zuletzt kraft der daran beteiligten Künstlichen Intelligenz wie ein intelligentes Subjekt verhalten sollte, ist nicht gesagt, dass dieses Verhalten unseren Zwecken dient und unsere Probleme löst³⁸.

Künstliche Intelligenz im Anthropozän lässt demnach den hier vorgelegten Begriff der Technik zu einem Problembezug werden: zum Begriff eines Problems, auf das wir als endliche Subjekte stoßen, auch weil es sich auf eine gewisse Weise anfühlt, sich in einer derartigen Situation zu befinden. Da der Begriff der Künstlichen Intelligenz demjenigen der Technik ohnehin implizit innewohnt, wird dadurch Technik im Anthropozän insgesamt zu einem Problem für endliche, menschliche Intelligenz. Wie eingangs angedeutet, könnte ein Begriff der Natur in dieser problematischen Situation weiterhelfen, was nun zu zeigen ist.

³⁴ Vgl. Schmidt/Zapf (2021).

³⁵ Zu Narrativen des Anthropozäns als Ausdruck dieses kulturellen Reflexes vgl. Schmaus (2020).

³⁶ Vgl. Mainzer (2020) sowie als kritische Reflexion zu diesem Thema Rathmann (2021). Chancen und Risiken eines entsprechenden Zusammenspiels natürlicher und künstlicher Intelligenz wägt aus psychologischer Perspektive Meitner (2021) ab.

³⁷ Vgl. Lovelock (2019).

³⁸ Zu diesem Problem vgl. Bostrom (2014), S. 152–154.

6. Ein Begriff der Natur

Subjekte im Anthropozän stehen also vor der Herausforderung, sich mit ihrem je eigenen Standpunkt auseinanderzusetzen, und dies in einer Umwelt, die so stark technisch überformt ist, dass sich die Frage stellt, ob diese Umwelt noch dem Begriff der Natur genügt, ja sogar, ob dieser Begriff überhaupt (noch) brauchbar ist³⁹. Damit schließt sich ein begriffsgeschichtlicher Kreis (und eröffnet, wie das üblich ist, sogleich einen neuen, weitergespannten), da in der abendländischen Tradition die Arbeit am Naturbegriff immer schon der Bestimmung des Standpunkts der Subjekte dient, die sie leisten. Diese Arbeit vollzieht sich in drei Schritten, die hier in einer verallgemeinernden Weiterführung von Einteilungen dargestellt werden, die sich bei Elisabeth List im Anschluss an Serge Moscovici⁴⁰ und bei Jens Soentgen im Anschluss an Gregor Schiemann⁴¹ finden. Jene Schritte führen von einem intrinsischen über einen extrinsischen hin zu einem relationalen Naturbegriff.

Die abendländische Arbeit am Naturbegriff beginnt aus gutem Grund in der Frühzeit der griechischen Philosophie, die sich mit der Erfahrung einer großen Fülle von Standpunkten konfrontiert sah: schon innerhalb der griechischen Stadtstaaten mit ihren vielfältigen politischen und kulturellen Verfasstheiten und erst recht im Kontakt mit unterschiedlichen Hochkulturen⁴². Angesichts dessen stellte sich die Frage, wie sich inmitten dieser Mannigfaltigkeit eine begründende und zugleich begründende Rede (ein Logos) rechtfertigen ließe⁴³. Ein Versuch, diese Frage zu beantworten, besteht darin, etwas zu finden, worauf von jedem Standpunkt aus jeweils der gleiche Bezug möglich ist. Was dabei gefunden wird, ist dasjenige, was allem Anschein nach jedem Standpunkt gleichsam zuwächst, weil es sich allenthalben von selbst entfaltet: die Natur⁴⁴. Dieser Begriff wird dabei auf eine zweifache Weise erläutert⁴⁵: vom Umfang her als Inbegriff der vertrauten natürlichen Arten, vom Inhalt her auf klassische Weise bei Aristoteles als Inbegriff all dessen, was das

³⁹ Vgl. McGrath (2019), Kap. 5.

⁴⁰ Vgl. List (2007), S. 165–168.

⁴¹ Vgl. Soentgen (2020), S. 116f.

⁴² Vgl. Habermas (2020), S. 417.

⁴³ Vgl. Zorn (2016), Kap. 3.

⁴⁴ Vgl. Buchheim (1994), S. 91–95, 152–154.

⁴⁵ Zum Folgenden vgl. Schiemann (2005), Kap. 1.1.1.

Prinzip seiner Bewegung und seines Ruhens in sich selbst trägt. Mit der so verstandenen Natur sind wir von unseren jeweiligen Standpunkten her vertraut, sofern wir selbst zum Natürlichen gehören, und können daher das Natürliche mit geläufigen Namen belegen, auch wenn wir uns dabei gelegentlich berichtigen (lassen) müssen. Auf Natur beziehen wir uns daher gleichsam von innen. Subjektivität wird dabei als etwas verstanden, was mit Natürlichen eng verflochten ist, wie es die aristotelische Formel von der Seele als der Form eines natürlichen, organischen Körpers zum Ausdruck bringt. Die Plausibilität dieses Naturbegriffs beruht darauf, dass er sich vom Standpunkt der endlichen, menschlichen Subjektivität aus anbietet, die sich ja aus sich selbst heraus entfaltet und sich dabei als in einer Welt sich ebenfalls, wenn auch gelegentlich auf anderer Weisen, entfaltender Wesenheiten beheimatet erfährt. Der intrinsische Naturbegriff ist daher sowohl anthropozentrisch (von einem menschlichen Standpunkt aus gedacht) als auch anthropomorph (nach Maßgabe menschlicher Standpunkthaftigkeit gedacht). Wie sich das Subjekt auf seinem Standpunkt als eine Einheit denkt, so denkt es mit dem intrinsischen Naturbegriff auch das Natürliche als eine Menge von Gegenständen. Im Lichte dieses Naturbegriffs erscheint die Welt als ein Kosmos, als eine schöne hierarchische Ordnung aus einzelnen Dingen, die sich kraft ihrer inneren Prinzipien (ihrer „Naturen“) darin einfügen⁴⁶.

Dieser Naturbegriff hat sich als äußerst einflussreich erwiesen; er wirkt bis in unsere heutige Rede von echter, unverfälschter Natur, von Wildnis, hinein⁴⁷. Dieser Naturbegriff ist es auch, der das Technische und damit auch Künstliche Intelligenz scharf vom Natürlichen abzugrenzen scheint, insofern das Technische das Prinzip seiner Bewegung und seines Ruhens eben nicht in sich selbst trägt, sondern von außen aufgeprägt erhalten hat.

Allerdings steht der intrinsische Naturbegriff vor einer Herausforderung: Mit ihm wird Natur gedacht als eine Mannigfaltigkeit möglicher Standpunkte von Subjektivität. Von welchem Standpunkt aus geschieht dies eigentlich, von was für einem Standpunkt aus wäre so etwas überhaupt möglich? Offenbar wird in diesem Denken doch ein Standpunkt gegenüber der Natur bezogen, jedenfalls gegenüber dem, was als Bewegtes und Ruhendes unter jenen Begriff fällt. Diese Herausforderung wurde bereits in der frühen griechischen Philosophie durch die Eleaten scharf formuliert.

⁴⁶ Vgl. Voigt/Voigt (2011).

⁴⁷ Vgl. Schiemann (2005), S. 31f.

Aristoteles versuchte, ihr dadurch zu begegnen, dass er zwar die Seele als Trägerin von Subjektivität als Form mit ihrem Körper verbindet, ihr aber im Fall des Menschen als einem denkenden Wesen noch einen diese Denkfähigkeit stiftenden Teil zuschreibt, den Geist (*nous*), der „von draußen kommt“, der die Natur eben deshalb auf den Begriff bringen kann, weil er ihr nicht (ganz) angehört⁴⁸.

Was mit diesem Geist gemeint ist, daran hat sich die aristotelische Tradition und mit ihr die Hauptströmung der Arbeit am intrinsischen Naturbegriff abgearbeitet und auch darüber ist sie in der frühen Neuzeit in eine Krise geraten⁴⁹. In dieser Krise macht René Descartes gleichsam aus der Not eine Tugend, indem er das Verhältnis zwischen Subjekt und Natur neu bestimmt⁵⁰: Natur ist auch für Descartes ein Bezugspunkt für Subjektivität. Nur denkt er sich Subjektivität nun nicht mehr inmitten des Natürlichen, sondern, kraft des von ihm angewandten methodischen Zweifels, streng ihm gegenüber. Subjektivität kann sich auf Natürliches beziehen, da sie Begriffe in sich vorfindet, mit denen sich dessen Bewegungen und Ruhezustände denken lassen, allen voran den Begriff der Ausdehnung; und Natürliches kann sich auf Subjektivität beziehen lassen, da es unter diese Begriffe fällt. Das Gegenüber von Subjektivität und Natur wird von Descartes auf eine bekannte Formel gebracht: diejenige des Gegensatzes zwischen dem Subjekt als denkendem Ding, das nicht ausgedehnt ist, und dem Natürlichen als ausgedehntem Ding, das nicht denkt. Da hierbei das Natürliche unter allgemeine Begriffe fällt, die sich mit mathematischer Präzision denken lassen, ist es als ein Bereich strenger Gesetzmäßigkeit zu verstehen, die sich mit Hilfe jener Begriffe formulieren lässt – als das Reich, in dem die Naturgesetze gelten. Diesem Bereich steht Subjektivität gegenüber, indem sie sich in freier Selbstreflexion auf ihr eigenes Wesen, das Denken, und das Wesen des von ihr gedachten Natürlichen besinnt. Damit ist vom Standpunkt der Subjektivität aus gesehen ein extrinsischer Naturbegriff gewonnen: Das Natürliche ist demgemäß zum einen dasjenige, was dem Subjekt gegenüber äußerlich ist, und zum anderen dasjenige, was auch in sich selbst als Mannigfaltigkeit von Äußerlichkeiten begriffen werden kann, als Körper, deren Verhältnisse zueinander nicht schon aus ihnen selbst heraus bestimmt sind, wie dies nach intrinsischem Verständnis der Fall war, sondern durch die Naturgesetze.

⁴⁸ Vgl. Verf. (2007).

⁴⁹ Vgl. Schiemann (2005), S. 165.

⁵⁰ Vgl. zum Folgenden ebd., Kap. 1.2.1. und 1.2.3.

Dieser extrinsische Naturbegriff ist nicht mehr anthropozentrisch, da er sich nicht mehr auf den Standpunkt des Menschen als eines (auch) natürlichen Wesen beruft; er ist vielmehr azentrisch, da es weder im mannigfaltigen Natürlichen noch im Verhältnis zwischen Subjekt und Natur ein Zentrum gibt; damit wird er auch dem wissenschaftsgeschichtlichen Aufbruch aus einem scheinbar geschlossenen Kosmos in ein Universum gerecht, das sich zumindest zunächst als unendlich zeigt⁵¹. Der extrinsische Naturbegriff ist auch nicht mehr anthropomorph, denn der Mensch, als Naturwesen verstanden, gilt nun nur noch als ein Körper unter beliebigen anderen. Die Gestaltung der Körper vollzieht sich vielmehr nach Maßgabe der Naturgesetze und ist daher als eine Konstruktion auf mathematischer Grundlage aufzufassen. Wenn Körper bestimmte komplexe Aktivitäten aufweisen, dann deshalb, weil es sich bei ihnen um entsprechend konstruierte Maschinen, um Automaten handelt. So erweist sich der extrinsische Naturbegriff als technomorph. Den Produkten der Technik steht allerdings von vornherein Subjektivität und damit auch Intelligenz im oben angegebenen Sinn gegenüber. Der Geist kommt in die Natur aus dieser Perspektive gesehen nicht nur von draußen, sondern er bleibt auch dort.

Auch dieser Naturbegriff hat eine mächtige Wirkungsgeschichte entfaltet und nicht zuletzt die technische Überformung der Umwelt im Anthropozän denkbar und machbar erscheinen lassen⁵². Aber auch der extrinsische Naturbegriff hat mit einer Herausforderung zu kämpfen, die sich in der Frage artikuliert, wie denn Subjekt und Natur angesichts ihrer grundständigen Verschiedenheit miteinander in Beziehung stehen können. Descartes selbst versucht diese Frage für sich selbst dadurch zu beantworten, dass er sich dabei wiederum auf Natur beruft, wobei er dann aber unter Natur dasjenige versteht, was das denkende Subjekt und sein ausgedehntes Gegenüber miteinander verbindet, und zwar im Fall des verkörperten Menschen, wie Descartes selbst zugesteht, auf eine sehr innige Art und Weise. Die Lehre, die aus dieser zunächst irritierenden Verwendung des Naturbegriffs gezogen werden kann, lautet: Die Gegenüberstellung von Subjekt und Natur lässt sich nur von einem Standpunkt her denken, der diese Gegenüberstellung im Sinne einer bloßen Verschiedenheit bereits überwunden hat, von dem her diese Verschiedenheit als eine wechselseitige Bezogenheit gedacht werden kann. Das denkende Subjekt hat sich bereits als eine derartige Bezogenheit von seinem Standpunkt her

⁵¹ Vgl. Koyré (1969).

⁵² Vgl. McGrath (2019), S. 156f.

erwiesen. Wechselseitigkeit lässt sich daher dadurch denken, dass auch dem Natürlichen jeweils eine Standpunkthaftigkeit zugeschrieben wird. So verstanden, ist Natur Inbegriff dessen, was sich von einem Standpunkt her auf anderes beziehen kann, das ebenfalls einen Standpunkt einnehmen kann. Charakteristisch für das Natürliche ist danach eben jene mögliche wechselseitige Bezogenheit, so dass wir hier von einem relationalen Naturbegriff reden können. Von diesem relationalen Naturbegriff her stellt sich das Natürliche als ein Geflecht von Beziehungen möglicher Standpunkte dar. Unabhängig von begriffsgeschichtlicher Reflexion hat dies Saul Kripkes Analyse der Logik erarbeitet, die mit der Benennung natürlicher Arten verbunden ist⁵³: Wir begegnen Exemplaren dieser Arten und nehmen in diesem Sinne eine Stichprobe. Bei diesem Anlass verleihen wir jenen Arten einen Namen, der sich als starrer Bezeichnungsausdruck auf sie bezieht, was immer diese Arten als solche ausmacht. Was diese Arten als solche ausmacht, ist entweder in der Stichprobe unmittelbar gegeben – wenn es sich um eine direkte Empfindungsqualität handelt – oder lässt sich eventuell durch weitere Untersuchungen feststellen, wie im Fall der geläufigen biologischen Arten, die Kripke hier als Beispiel heranzieht. Dieses Beispiel mag wiederum anthropozentrisch und anthropomorph erscheinen, denn es sind ja jeweils menschliche Instanzen, die derartige Benennungen und Untersuchungen vornehmen. Doch ist es nicht nur denkbar, sondern auch erfahrbare, dass auch Menschen ihrerseits gleichsam zum Gegenstand einer Stichprobe werden, wie dies Val Plumwood nach ihrer lebensbedrohlichen Begegnung mit einem Krokodil herausgearbeitet hat⁵⁴. Auch auf Menschen kann Bezug genommen werden, wenngleich wir dies, wie Plumwood meint, im Rahmen unserer westlichen Kultur gerne verdrängen. Auch hier herrscht also zumindest prinzipiell Wechselseitigkeit.

Diesen relationalen Naturbegriff haben List und Soentgen auf zwei unterschiedliche und gerade darin charakteristische Weisen bezeichnet: Für List handelt es sich dabei um den Begriff eines „kybernetischen Naturzustands“⁵⁵; Soentgen bemüht sich dagegen um einen „ökologischen“ Naturbegriff⁵⁶. Im ersten Fall wird die Verbindung, die für den Naturbegriff zentral ist, von der Technik und damit vom ext-

⁵³ Vgl. Kripke (1981), S. 133–147, 153–164, 178f., 184; ders. (2014), S. 70–83.

⁵⁴ Vgl. Plumwood (2012), S. 9–21.

⁵⁵ List (2007), Kap. IX–X.

⁵⁶ Soentgen (2020), S. 116–118.

rinsischen Naturbegriff her gedacht, oder, wie wir jetzt sagen können, vom extrinsischen Pol des relationalen Naturbegriffs; im zweiten Fall wird jene Verbindung vom Naturwüchsigen und damit auch noch vom intrinsischen Naturbegriff beziehungsweise vom intrinsischen Pol des relationalen Naturbegriffs her gedacht. Von der Warte des relationalen Naturbegriffs her stellt sich das Natürliche als Inbegriff von Beziehungen dar, die jeweils diesen „bipolaren“ Charakter aufweisen, also jeweils schon den eigenen subjektiven Standpunkt als Gegenüber eines anderen ebenfalls subjektiven Standpunkts auffassen lassen⁵⁷. Der relationale Naturbegriff ist demnach polyzentrisch.

Insofern dabei das Gegenüber als Standpunkt eines Subjekts gedacht wird, kann jenes Subjekt und mithin sein Standpunkt ebenso wenig unter eine rein äußere Bestimmung fallen wie das Subjekt, von dem aus dies jeweils gedacht wird. Damit zeigt sich die reflexionslogische Grundlage⁵⁸ der Beobachtung, die Kripke im Hinblick auf das Benennen natürlicher Arten angestellt hat: Benennung muss nicht mit einer hinreichenden Bestimmung verbunden sein, lässt diese vielmehr zumindest zunächst noch offen. Das Natürliche im Sinne des relationalen Naturbegriffs ist demnach der Bereich dessen, was wir benennen können, ohne dass wir es dafür auch schon hinreichend bestimmen, etwa umfassend beschreiben, müssen. Nachträgliche Beschreibungen und ihnen entsprechende Gestaltungen werden dadurch nicht ausgeschlossen, aber auch nicht vorbestimmt; daher ist der relationale Naturbegriff gewissermaßen polymorph.

Der relationale Naturbegriff passt damit gut zu einem Verständnis der neueren Wissenschaften, dem zufolge diese es von ihrem Gegenstandsbereich mit einem Geflecht von Beziehungen zu tun haben⁵⁹ und auch – gerade im Anthropozän – vor der Aufgabe stehen, sich interdisziplinär untereinander zu vernetzen⁶⁰.

Auch der relationale Naturbegriff steht vor einem Problem, nämlich der Frage, wie sich der je eigene Standpunkt von außen denken lässt und wie sich andere

⁵⁷ Diese These lässt sich durch die begründete Annahme unterstützen, dass jeder mögliche Standpunkt der Standpunkt eines umgreifenden, transzendentalen Subjekts ist. Vgl. Meixner (2017), v. a. S. 396–398.

⁵⁸ Zur Reflexionslogik vgl. Zorn (2016), Kap. 2.

⁵⁹ Vgl. Cassirer (2000).

⁶⁰ Vgl. Soentgen (2021); St. Voigt (2021).

Standpunkte jeweils von ihrer eigenen subjektiven Innenseite her denken lassen⁶¹. Wie sich im Lauf dieser Überlegungen gezeigt haben dürfte, ist der Naturbegriff aber ohnehin ein Problembegriff, nämlich ein Begriff davon – oder zumindest ein Versuch, zu begreifen –, wie sich ein Subjekt jeweils sowohl in Bezug zu sich selbst als auch in Bezug zu anderen Subjekten denken kann. Ein Vorzug des relationalen Naturbegriffs besteht darin, diese Problematik nicht etwa zu verdecken, sondern sie in ihrer Komplexität weiter zu entfalten.

7. Natürlichkeit und Künstlichkeit von Intelligenz im Anthropozän

Der Blick auf einen relationalen Naturbegriff gestattet eine Antwort auf die eingangs gestellte Frage, ob und wie Künstliche Intelligenz im Anthropozän auch natürlich sein kann, ja sogar sein soll. Dem relationalen Naturbegriff zufolge ist natürlich, was sich in einen Wechselbezug einfügt, in dem ein Subjekt seinen eigenen Standpunkt als das Gegenüber eines anderen Subjekts denkt und dieses Gegenüber wiederum als den möglichen Standpunkt eines anderen Subjekts anerkennt. So gesehen ist Intelligenz allgemein dann natürlich, wenn sie sich als reflektiertes Problembewusstsein gestaltet und dadurch die eigene Standpunkthaftigkeit sowie die qualitative Eigenständigkeit des jeweiligen Problems erkennt. Technik kann, wie wir gesehen haben, diesen Zusammenhang überbrücken, indem sie das Problem zu lösen versucht, anstatt es als Anlass zur Reflexion zu nutzen, und indem das reflexionsfähige Subjekt in ihrer Anwendung etwas anderes als sich selbst zum Zweck setzt. Technik, die nichts anderes tut, als jenen Zusammenhang zu überbrücken und damit auszublenken, steht durchaus im Gegensatz zur (relational verstandenen) Natur, weil dies die Reflexion auf die Zusammenhänge behindert, in der das diese Technik anwendende Subjekt jeweils steht, und wäre in einem schlechten Sinne künstlich. Technik kann das Subjekt gerade dadurch aber auch entlasten, da es sich dank ihr nicht mit beliebigen Problemen befassen muss und dabei seine Reflexionsfähigkeit durchaus

⁶¹ Daran (unter dem Titel „Kombinationsproblem“) arbeitet sich der zeitgenössische Panpsychismus ab, der als diejenige Fassung der Philosophie des Geistes verstanden werden kann, die am besten zu einem relationalen Naturverständnis passt. Vgl. dazu Verf. (2020c).

bewahren kann. Solcherart gestaltete Technik kann geradezu in den Dienst subjektiver Selbstreflexion treten⁶². Als technisch herbeigeführt steht starke Künstliche Intelligenz – und allein von ihr lohnt es sich hier zu handeln, wie es sich erwiesen hat – vor derselben Problematik, die im Anthropozän als einer planetenumspannend technisch überformten Umwelt auf die Spitze getrieben worden ist. Diese Problematik lässt sich nicht mehr rein technisch bewältigen, da sie sich dadurch nur perpetuieren würde. Intelligenz im Anthropozän – menschliche, Künstliche und auch sonstige außermenschliche – sieht sich damit vor die Aufgabe gestellt, ihre Reflexionsfähigkeit zu bewahren und wenn möglich zu steigern, um der Komplexität der Situation gerecht zu werden. Einen Maßstab dafür kann das Ausmaß darstellen, indem es dabei gelingt, sich als Intelligenz in die wechselseitigen Bezüge einzufügen, die auch und vielleicht gerade eine technisch überformte Umwelt nach wie vor bietet, und in diesem Sinne natürlich zu sein oder sogar zu werden. Wie auch immer geartete Künstliche Intelligenz wird jeglicher natürlichen Intelligenz innerhalb dieser Umwelt ohnehin im Rahmen der sie prägenden Technosphäre begegnen. Ob und wie es gelingt, diese wechselseitigen Bezüge als einen Erfahrungsraum gemeinsam gelebter Reflexion und damit auch Vernünftigkeit zu erhalten, in diesem Sinne zu kultivieren⁶³ und weiterzuentwickeln⁶⁴, mag ein Gütekriterium und damit eine lohnenswerte Herausforderung für Intelligenz aller Art im Anthropozän darstellen.

Literatur

- Beck, Heinrich (²1979): *Kulturphilosophie der Technik. Perspektiven zu Technik – Menschheit – Zukunft*, Trier
- Belke, Annette (2021): Facetten natürlicher Intelligenz am Beispiel des Brown Bear/Grizzlybär (*Ursus arctos horribilis*), in *vorliegendem Band*, S. 293–310
- Borner, Marc (2016): *Über präreflexives Selbstbewusstsein. Subpersonale Bedingungen – empirische Gründe*, Münster

⁶² Dies zeigt, am Beispiel der frühneuzeitlichen Technikverwendung bei da Vinci und Galilei, Heichele (2016).

⁶³ Zu dieser Aufgabe der Gestaltung eines Begegnungsgeschehens von Natur, menschlichem Geist und Technik vgl. Beck (²1979).

⁶⁴ Zu diesen Anliegen vgl. Friedrich/Rathmann (2021) und Meixner (2020).

- Bostrom, Nick (2014): *Superintelligenz. Szenarien einer kommenden Revolution*. Aus dem Englischen von Jan-Erik Strasser, Berlin
- Buchheim, Thomas (1994): *Die Vorsokratiker. Ein philosophisches Porträt*, München
- Cassirer, Ernst (2000): *Substanzbegriff und Funktionsbegriff*, Darmstadt (Gesammelte Werke. Hamburger Ausgabe, Bd. 6)
- Claessens, Dieter (2019): Art. „Entlastung“, in: *Historisches Wörterbuch der Philosophie*, Bd. 2 D-F, Sonderausgabe Darmstadt 2019, Sp. 538f.
- Crutzen, Paul J. (2019): *Das Anthropozän. Schlüsseltexte des Nobelpreisträgers für das neue Erdzeitalter*. Hg. von Michael Müller, München
- Davis, Martin (2011): *The Universal Computer. The Road from Leibniz to Turing*, Boca Raton–London–New York
- Dörner, Dietrich (2021): Mülltonne, Speerschleuer und Fahrradschlauch – Über künstliche und natürliche Intelligenz, in *vorliegendem Band*, S. 217–233
- Dufresne, Todd (2019): *The Democracy of Suffering. Life on the Edge of Catastrophe, Philosophy in the Anthropocene*, Montreal–Kingston
- Fichte, Johann Gottlieb (1794/2017): *Grundlage der gesamten Wissenschaftslehre als Handschrift für seine Zuhörer (1794)*. Mit einer Einleitung und Registern hg. von Wilhelm G. Jacobs. Nachdruck der 4. Auflage 1997, Hamburg 2017
- Friedrich, Marion (2021): Intelligenz aus philosophisch-psychologischer Sicht, in *vorliegendem Band*, S. 135–162
- Friedrich, Marion/Rathmann, Joachim (2021): Corona und die Herausforderung für den Umweltschutz, in *vorliegendem Band*, S. 235–252
- Habermas, Jürgen (*2020): *Auch eine Geschichte der Philosophie. Bd. 1: Die okzidentale Konstellation von Glauben und Wissen*, Frankfurt am Main
- Heichele, Thomas (2016): *Die erkenntnistheoretische Rolle der Technik bei Leonardo da Vinci und Galileo Galilei im ideengeschichtlichen Kontext*, Münster
- , Hg. (2020): *Mensch – Natur – Technik. Philosophie für das Anthropozän*, Münster
- (2021): Künstliche Intelligenz im Lichte der Technikphilosophie, in *vorliegendem Band*, S. 79–108
- Heidegger, Martin (1953/2000): Die Frage nach der Technik, in: ders., *Vorträge und Aufsätze*, Frankfurt am Main, S. 5–36 (Gesamtausgabe, 1. Abt., Bd. 7)
- Henrich, Dieter (2019): *Dies Ich, das viel besagt. Fichtes Einsicht nachdenken*, Frankfurt am Main
- Horn, Eva/Bergthaller, Hannes (2019): *Anthropozän zur Einführung*, Hamburg
- Höslé, Vittorio (2018): *Kritik der verstehenden Vernunft. Eine Grundlegung der Geisteswissenschaften*, München

- Irrgang, Bernhard (2020): *Roboterbewusstsein, automatisiertes Entscheiden und Transhumanismus. Anthropomorphisierungen von KI im Licht evolutionär-phänomenologischer Leib-Anthropologie*, Würzburg
- Klingan, Katrin/Rosol, Christoph, Hg. (2019): *Technosphäre*, Berlin
- Koyré, Alexandre (1969): *Von der geschlossenen Welt zum unendlichen Universum*. Übersetzt von Rolf Dornbacher, Frankfurt am Main
- Kripke, Saul A. (1981): *Name und Notwendigkeit*. Übersetzt von Ursula Wolf, Frankfurt am Main
- (2014): *Referenz und Existenz. Die John-Locke-Vorlesungen*. Aus dem Englischen übersetzt von Uwe Voigt, Stuttgart
- Latour, Bruno (2017): *Kampf um Gaia. Acht Vorträge über das neue Klimaregime*. Aus dem Französischen von Achim Russer und Bernd Schwibs, Frankfurt am Main
- List, Elisabeth (2007): *Vom Darstellen zum Herstellen. Eine Kulturgeschichte der Naturwissenschaften*, Weilerswist
- Lovelock, James (2016): *Gaia. A new look at life on Earth*, Oxford
- (2019): *Novacene. The Coming Age of Hyperintelligence*, London
- Mainzer, Klaus (1995): *Computer – neue Flügel des Geistes?*, Berlin
- (2020): Vom Anthropozän zur Künstlichen Intelligenz. Herausforderungen von Mensch und Natur durch Technik im 21. Jahrhundert, in: Thomas Heichele, Hg., *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster, S. 155–168
- McGrath, Sean (2019): *Thinking Nature. An Essay in Negative Ecology*, Edinburgh
- (2021): AI and the Human Difference, in *vorliegendem Band*, S. 53–77
- Meitner, Michael J. (2021): Artificial Intelligence: Thoughts from a Psychologist, in *vorliegendem Band*, S. 163–176
- Meixner, Uwe (2017): Idealism and Panpsychism, in: Brüntrup, Godehard/Jaskolla, Ludwig, Hg., *Panpsychism. Contemporary Perspectives*, Oxford
- (2020): Natur und Vernunft im Anthropozän, Thomas Heichele, Hg., *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster, S. 67–84
- (2021): Bewusstseinsintelligenz und Künstliche Intelligenz, in *vorliegendem Band*, S. 13–31
- Pape, Helmut (1997): *Die Unsichtbarkeit der Welt. Eine visuelle Kritik neuzeitlicher Ontologie*, Frankfurt am Main
- (2006): Kontrollierte Abstraktion und Selbstkritik, in: Wolfgang Achtnr u. a., Hg., *Künstliche Intelligenz und menschliche Person*, Marburg, S. 107–121
- Plumwood, Val (2012): *The Eye of the Crocodile. With an introduction by Deborah Bird Rose, Freya Matthews and Kate Rigby*. Ed. by Lorraine Shannon, Canberra
- Popitz, Heinrich (1995), *Der Aufbruch zur Artifizialen Gesellschaft. Zur Anthropologie der Technik*, Tübingen

- (1997): *Wege der Kreativität*, Tübingen
- Rathmann, Joachim (2021): Künstliche Intelligenz im Umweltschutz: Möglichkeiten und Grenzen, in *vorliegendem Band*, S. 253–275
- Renn, Jürgen/Scherer, Bernd, Hg. (2017): *Das Anthropozän. Zum Stand der Dinge*, Berlin
- Rosengrün, Sebastian (2021a): Was ist KI und wenn ja, wie viele? Vier Rätsel einer Philosophie der Künstlichen Intelligenz, in *vorliegendem Band*, S. 33–52
- (2021b): *Künstliche Intelligenz zur Einführung*, Hamburg
- Schiemann, Gregor (2005): *Natur, Technik, Geist. Kontexte der Natur nach Aristoteles und Descartes in lebensweltlicher und subjektiver Erfahrung*, Berlin–New York
- Schmaus, Thomas (2020): „Erzähl uns deine Erdgeschichte!“ Narrative Identität im Anthropozän, in: *Comenius-Jahrbuch* 28, S. 33–54
- Schmidt, Matthias/Zapf, Hubert, Hg. (2021): *Environmental Humanities*, Göttingen
- Sloterdijk, Peter (1989): *Eurotaoismus. Zur Kritik der politischen Kinetik*, Frankfurt am Main
- Soentgen, Jens (2014): Der Lech als Cyborg, in: Marita Kraus u. a., Hg., *Der gezähmte Lech. Fluss der Extreme*, München, S. 151–160
- (2018): *Ökologie der Angst*, Berlin
- (2019): *Konfliktstoffe. Über Kohlendioxid, Heroin und andere strittige Substanzen*, München
- (2020): Der ökologische Naturbegriff, in: Thomas Heichele, Hg., *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster, S. 115–130
- (2021): „Wer nichts als Chemie versteht, versteht auch die nicht recht.“ Interdisziplinarität als Mehrsprachigkeit, in *vorliegendem Band*, S. 277–292
- Voigt, Stefanie (2021): Warum Data malt – Interdisziplinarität und Ästhetik, in *vorliegendem Band*, S. 199–213
- Voigt, Stefanie/Voigt, Uwe (2011): Head jewellery – a theory of the theory of jewellery, in: Wilhelm Lindemann, Hg., *Thinking Jewellery: on the way towards a theory of jewellery*, Stuttgart, S. 80–93
- Voigt, Uwe (2007): Wozu braucht Aristoteles den „Geist von draußen“ in seinen biologischen Schriften?, in: *AKAN – Antike Naturwissenschaft und ihre Rezeption* 17, S. 29–38
- (2020a): Das Anthropozän als geistige Umweltkrise, in: Thomas Heichele, Hg., *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster, S. 85–102
- (2020b): Wissen um Atmosphären – Bildung für das Anthropozän?, in: *Comenius-Jahrbuch* 28, S. 13–32
- (2020c): Eingestimmte Subjekte? Das Kombinationsproblem des Panpsychismus im Licht der Atmosphärenkonzeption der Neuen Phänomenologie, in: Wolf, Barbara/Julmi, Christian, Hg., *Die Macht der Atmosphären*, Freiburg im Breisgau, S. 60–74
- Zorn, Daniel-Pascal (2016): *Vom Gebäude zum Gerüst. Entwurf einer Komparatistik reflexiver Figuren in der Philosophie*, Berlin

Psychologische Perspektiven

Intelligenz aus philosophisch-psychologischer Sicht

Marion Friedrich

Zusammenfassung

Unter den Begriff der Intelligenz werden alle intelligenten Systeme subsumiert. Hier wird die These aufgestellt, dass alle intelligenten Systeme natürlicher Intelligenz entsprechen, während Menschliche Intelligenz (MI) mit anderen animalischen Intelligenzen koexistiert und ihrerseits Künstliche Intelligenz (KI) a) hervorgebracht und b) als Instrument/Werkzeug nutzt. So ist im Rahmen von KI besser von deren Funktionalität statt von Intelligenz zu sprechen.

Bislang in der Psychologie diskutierte Konzepte menschlicher Intelligenz können nicht auf diese Funktionalität künstlicher Systeme angewendet werden. Auch MI fordert im Anthropozän eine neue Betrachtungsweise und sollte um ein ethisch-pragmatisches Verständnis ergänzt werden.

Des Weiteren stellt sich die Frage, ob im Zeitalter des Anthropozäns überhaupt noch von menschlicher Intelligenz als Fähigkeit der Problemlösung gesprochen werden kann. Angelehnt an Wittgensteins Sprachspieltheorie ist es vielleicht unsinnig, zu fragen, was Intelligenz ist – wichtig ist, was man damit tut. Die Finalität natürlicher wie künstlicher Intelligenz muss untersucht werden und dabei von Intentionalität/Volition unterschieden werden.

Menschliche Intelligenz (MI) oder HI (Human Intelligence) wird hierbei als Teilbereich der Natürlichen Intelligenz (NI) aufgefasst und evolutionär als Resultat der Anpassungsnotwendigkeit und -fähigkeit auf sich ändernde Umweltbegebenheiten verstanden. Dieses Verständnis lässt die Weiterführung des Intelligenzbegriffs als Entwicklungsmöglichkeit neuer Fähig- und Fertigkeiten zu.

Künstliche Intelligenz (auch AI als Artificial Intelligence) hingegen wird verstanden als die von einer NI zielgerichtet entwickelte Funktionalität eines Werkzeugs,

d. i. in ihrer Ausprägung messbar an dem Grad der Aufgabenerfüllung. Somit ist KI MI unterzuordnen. Ein anthropologischer Fehlschluss wäre, KI aufgrund der (programmierten) Finalität Intentionalität zu unterstellen.

In Bezug auf MI ist deren Finalität jedoch unbedingt mit Intentionalität/Volition verbunden und richtet sich damit implizit oder explizit nach ethischen Normen aus. So ist mensch als fully ethical agent¹ zu verstehen (und auch in die Pflicht zu nehmen!).

MI muss unter dem finalitätsgerichteten Aspekt von Intentionalität und Volition untersucht werden: Der Mensch nutzt seine (und die von ihm entwickelte künstliche) Intelligenz, um zu überleben bzw. um das Leben so angenehm wie möglich zu gestalten – KI hingegen beabsichtigt und will nicht sowie nichts.

Zeigt sich aus diesem Verständnis heraus der Mensch im Anthropozän als intelligent? Der Mensch passt nicht mehr sich und seine Handlungen an die Umwelt an, sondern er verändert die Umwelt derart, dass er selbst keine neuen Fähigkeiten mehr entwickeln und bereits erworbene nicht mehr nutzen muss (z. B. Lesen, Schreiben, Raumorientierung, soziale Fähigkeiten etc.).

Auf der anderen Seite beraubt er sich (aus hedonistischen Motiven heraus?) durch die fortlaufende Beeinträchtigung und Beeinflussung natürlicher Prozesse seiner Lebensgrundlage Natur. Aus dieser Perspektive heraus ist das Verständnis von Intelligenz als Fähigkeit zur Lösung neuer Probleme (nach Stern) nicht mehr gültig.

Anstatt sich nun auf die Lösung der hausgemachten Problematiken zu konzentrieren (z. B. könnte KI auch z. B. als APP dahin gehend genutzt werden, dass der ökologische Fußabdruck von Waren als Kaufentscheidungshilfe berechnet werden kann) müht sich der Mensch an Spekulationen über die Eventualität zukünftiger Bedrohung durch die von ihm geschaffene KI ab. Er befürchtet eine Überwältigung durch autonome KI, die dem Kontrollverlust im Zauberlehrling gleicht. Hier zeigt sich ein anthropologischer Gedankenfehler: KI ist in Intentionalität und ihrer Finalität eben nicht mit MI zu vergleichen und ist intentions- wie volitionsfrei. Ebenso wenig bewertet KI von sich aus (Über)Leben als gut und/oder Untergang/Tod als schlecht. Von KI ist also weder Unterstützung noch eine Gefährdung hinsichtlich des Fortbestands der Spezies Mensch zu erwarten.

Der Intelligenzbegriff sollte unter Berücksichtigung einer handlungsorientierten Ethik spezifiziert bzw. neu definiert werden. Bei künstlichen Systemen als Werk-

¹ Vgl. Moor (2006).

zeug/Instrument ist auf den Begriff der Intelligenz zugunsten des Begriffs der Funktionalität zu verzichten.

In der Definition von MI muss deren Intentionalität und damit moralische Ausrichtung im Sinne des (Überlebens-)Schutzes einer jeder Form von Intelligenz (auch natürliche und animalische Intelligenz) stärker gewichtet werden.

Abstract

translated by Kevin Fencil

At present, all types of intelligent systems are subsumed under the general term “intelligence.” It is the author’s view that all intelligence systems are forms of natural intelligence. Further, that human intelligence (HI) coexists alongside other forms of animal intelligence, both a) giving rise to artificial intelligence (AI) and b) utilizing it as an instrument or tool. As such, it is better to speak of AI in terms of its functionality rather than of intelligence.

The concepts of human intelligence currently in use in the field of psychology cannot be said to apply to the functionality of artificial systems. In the age of the Anthropocene, a new view of human intelligence which also takes into account an ethic-pragmatic understanding of the term is required. Furthermore, it is fair to ask whether continuing to see human intelligence as the ability to solve problems (see W. Stern) is justified at all.

According to H. Eysenck, intelligence may be classified alongside temperament, character and body type as variables of personality – and thus as a primarily physiological factor. When viewed through the lens of Wittgenstein’s theory of language games, however, it may be unnecessary to ask what intelligence is at all – when what is important is what one does with it. The finality of both natural and artificial intelligence must be examined and differentiated from the concepts of intentionality and volition what means what someone tries to achieve by using intelligence/certain methods.

I consider human intelligence (HI) to be a subset of natural intelligence (NI). Seen in evolutionary terms, HI can be viewed as the result of the necessity and ability to adapt to changing environmental circumstances. This in turn allows intelligence to be further understood as the possibility of developing new abilities and skills that can be employed in specific ways to improve the circumstances of life and the chance of survival, among other things.

By contrast, artificial intelligence can be understood to be the functionality of a tool developed for a specific use by a form of NI and thus measurable according to how well it fulfills its task. In other words, AI is a subset of HI.

It would be an anthropological mistake to ascribe intentionality to AI based on its (programmed) finality. Within the context of HI, however, it is absolutely essential that this finality be seen in connection with intention/volution/agency and therefore subject, either explicitly or implicitly, to ethical norms. HI is capable of “full ethical agency”, whereas AI is capable of “explicit ethical agency” at best².

The examination of the ethically influenced aspects of intentionality and volition allows for a clearer distinction between human intelligence and the functionality (utility, functioning) of a machine: Man uses his intelligence, and the artificial intelligence he has developed, in order to ensure survival (ideally not only his own or that of his species, but of all living beings) or in order to make life as pleasant as possible. AI, on the other hand, does not intend, much less want, anything. AI does not seek to ensure its own survival or the survival or elimination of other species.

Framed in this way, can mankind in the age of the Anthropocene be seen as intelligent? Man no longer adapts himself or his actions to the environment, but rather alters the environment to such a degree that he no longer has to develop new abilities or use the abilities he has already acquired (for example, reading, writing, spatial orientation, social skills). At the same time (for hedonistic reasons?), he robs himself of the nature upon which he depends for his existence by continually influencing and interfering with natural processes. It also appears as if routine use of manmade forms of AI negatively impact certain aspects of human intelligence (for example a decrease in the ability to spatially orientate and navigate due to navigation devices) and have a clearly negative impact on emotional intelligence (as measured by EQ).

Further examples include an assumed decrease in the ability to empathize (accompanied by an increase in so-called disorganized attachment typologies and dissociative disorders), as well as an increase in neurotic disorders (impulse control, fear and anxiety, and other behavioral disorders) in connection with increased media consumption and the use of smartphones.

Instead of searching for solutions to these problems, which are of our making (for example AI could be used to create an app that measures the environmental foot-

² Cf. Moor (2006).

print of goods and services to aid in making better purchase decisions), we occupy ourselves with speculation about future threats that could arise from the AI we've created. Our fear of being overwhelmed by autonomous AI resembles the loss of control in Goethe's *The Sorcerer's Apprentice*.

Here, an anthropological error in reasoning becomes evident: AI cannot be compared to HI in either its intentionality or its finality. It is free of both intent and volition. Nor does AI by itself judge life/survival to be good or demise/death to be bad. It follows that AI cannot be expected to either support or endanger the continued survival of the human species.

To sum up, I advocate the use of a new, ethic-pragmatic definition of intelligence. For artificial systems used as tools or instruments, the term "intelligence" should be rejected in favor of the term "functionality."

In defining human intelligence, more emphasis should be placed on the intentionality and the moral compass of that intelligence as it relates to protecting (the survival of) each and every form of intelligence, including natural and animal intelligence. Nobody (least of all AI) can save the world except ourselves – by using our own natural intelligence.

1. Was ist unter Intelligenz zu verstehen?

Nicht erst nachdem 2015 das Future of Life Institute zahlreiche Projekte subventionierte, die auf die Sicherstellung einer „für die Menschen vorteilhaften“ KI³ abzielen, und seit Stephen Hawking's zeitgleicher dystopischer Aussage, dass AI der menschliche Gattung ein Ende setzen könnte⁴, beunruhigt die Horrorvision von bewussten und bössartigen Robotern, die die Menschheit unterwerfen und/oder ausrotten wollen. Golems und Zauberlehrlinge spuken als zeitlose Schreckgespenster durch die Buchhandlungen, Hollywood hat schon längst den Marktwert solcher futuristischer Untergangsszenarien entdeckt. Doch bevor überhaupt über die Wahrscheinlichkeit von die Welt beherrschenden Superintelligenzen spekuliert werden kann, muss zu-

³ Siehe <https://www.vox.com/2015/7/2/11564066/elon-musk-backed-group-attempts-to-avert-judgment-day-with-ai-rules>, Zugriff: 17.10.2020.

⁴ Siehe <https://www.computerworld.com/article/2854997/stephen-hawking-says-ai-could-end-human-race.html>, Zugriff: 17.10.2020.

nächst geklärt werden, was unter dem Begriff „Intelligenz“ in diesem Zusammenhang verstanden wird und welche Rolle Intention und Volition hierbei spielen.

Es gilt im ersten Schritt, den oftmals unreflektiert verwendeten Begriff „Intelligenz“ zu betrachten und gegebenenfalls neu zu definieren. Kann im Zeitalter des Anthropozäns, in dem mensch befürchtet, dass das Menschenzeitalter bald vom „Robotozän“ abgelöst wird, der traditionell in der Psychologie gebrauchte Begriff „Intelligenz“ heute überhaupt noch Verwendung finden? Oder muss eine Neubestimmung erfolgen?

Kurz: Kann der traditionell-psychologische Begriff Intelligenz heute überhaupt noch Verwendung finden kann, ohne dass eine Neubestimmung vorangeht? Der Duden definiert Intelligenz als die „Fähigkeit, abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten“⁵. Das 1912 formulierte und bis heute verbreitete Verständnis der Menschlichen „Intelligenz als Anpassungsfähigkeit an sich ändernde Umweltbedingungen“ ist auf William Stern zurückzuführen⁶. Ähnlich definieren Kaplan und Haenlein Künstliche Intelligenz im Jahre 2010 als „a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation“⁷, also die Fähigkeit eines Systems, externe Daten korrekt zu interpretieren, aus diesen Daten zu lernen und diese Lernerfahrungen zu nutzen, um spezifische Ziele und Aufgaben durch flexible Anpassung zu erreichen.

Diese Entsprechung der Definitionen von MI und KI ist auch darauf zurückzuführen, dass der Mensch, der nicht über sich selbst bzw. die Grenzen seines Intellekts hinaus zu denken vermag, seine eigenen Intelligenzleistungen in KI zu reproduzieren versucht. Es ist nicht vorstellbar, was jenseits des eigenen Verständnisses (das Teil der Intelligenz ist) zu liegen vermag.

Schon in Bezug auf Natürliche (Menschliche) Intelligenz wirkt es beinahe unmöglich und sehr gewollt, die bisher geltenden Begrifflichkeiten in Relation zum Verhalten des „modernen“ Menschen im Anthropozän zu setzen – ist Verhalten überhaupt noch als intelligent zu bezeichnen, das natürliche Ressourcen mehr und mehr vernichtet, sich selbst die Lebensgrundlage raubt und eigene, angeborenen

⁵ Siehe <https://www.duden.de/rechtschreibung/Intelligenz>, Zugriff: 17.10.2020.

⁶ Vgl. Stern (1912).

⁷ Siehe <https://www.sciencedirect.com/science/article/pii/S0007681318301393?via%3Dihub>, Zugriff: 17.10.2020.

Fähigkeiten (von Empathievermögen bis hin zu Orientierungssinn) sukzessive durch Vernachlässigung verliert? Diese Art des Handelns erinnert doch sehr an die Osterinselbewohner, die der Überlieferung nach im 11.–16. Jahrhundert zum Transport der Moais alle Bäume abholzten und so die Erosion die Landschaft in Brachland verwandelte.

Hier kann der Begriff der Intelligenz nur dann noch Sinn ergeben, wenn er als Abstraktum, jenseits aller Handlungen und deren Konsequenzen, verwendet wird. Oder ist es eben als besonders intelligent zu betrachten, ganz im Sinne William Sterns, auch kognitive Leistungen mehr und mehr an technische und technologische Hilfsmittel zu delegieren, selbst, wenn diese Aufgabenübertragung schließlich kognitive Leistungseinbußen im Sinne der Fähigkeit, entsprechende Probleme zu lösen, mit sich bringt?

Wenn sich nur Teilbereiche des menschlichen Verhaltens als intelligent, hier verstanden als handlungs- und zielorientiert unter Berücksichtigung der natürlichen und sozialen Ressourcen auffassen lassen, so ist bei der Funktionalität Künstlicher Intelligenzen die Bezeichnung Intelligenz kontraintuitiv und kontraproduktiv. Ein Verständnis der Verhaltensregeln von Künstlichen Systemen als intelligent löst eine Flut von Projektionsfehlern aus, die bis zur Zuschreibung von Intentionalität, einer KI reicht. So werden dystopische Zukunftsszenarien befeuert und in der weiteren Übersteigerung Moravec'scher Eskalationen personifizierten Software-Agenten sogar Vernichtungsabsichten der Menschheit und/oder Streben nach der Weltherrschaft zu unterstellt⁸.

In diesem Beitrag wird die These aufgestellt, dass alle intelligenten Systeme Natürlicher Intelligenz entsprechen, wobei Menschliche Intelligenz mit anderen animalischen Intelligenzen koexistiert und ihrerseits Künstliche Intelligenz (KI)

- a. hervorgebracht hat und
- b. als Instrument/Werkzeug nutzt.

So ist im Rahmen von KI besser von deren Funktionalität statt von Intelligenz zu sprechen. MI (Menschliche Intelligenz) setzt sich dabei aus einzelnen Teilbereichen zusammen, ist aber als Gesamtfähigkeit nicht auf diese zu reduzieren. In ihren Teilbereichen wie Rechengeschwindigkeit, Merkfähigkeit (Speicherkapazität) etc. wird MI inzwischen häufig von KI übertroffen. Wenn jedoch auch der soziale Aspekt

⁸ Vgl. Moravec (1999).

und damit Emotionale Intelligenz wie auch die daraus resultierende Positionierung des Menschen als „wert-volles“ (nach Werten handelndes, ethisches Wesen) in Betracht gezogen werden, scheint eben dies ein relevantes Abgrenzungskriterium gegenüber Künstlichen Systemen zu sein.

MI kann dann so als Teil der Persönlichkeit, die ebenso Charakter, Temperament und biophysiological Merkmale – nach Eysenck (1916–1997): den „Körperbau“⁹ – umfasst, in ein übergeordnetes Konzept eingebettet werden. Dies fordert eine multifaktorielle Betrachtungsweise des Konzepts von Intelligenz.

Bislang in der Psychologie diskutierte bzw. bestehende Konzepte menschlicher Intelligenz können nicht auf diese Funktionalität künstlicher Systeme angewendet werden. Auch MI fordert im Anthropozän eine neue Betrachtungsweise und sollte um ein ethisch-pragmatisches Verständnis ergänzt werden. Der Mensch als zu intelligenten Handlungen fähig ist als *full ethical agent* nach Moor zu betrachten – KI hierbei ist als implizit oder maximal ansatzweise explizit ethischer Agent zu verstehen¹⁰.

Des Weiteren stellt sich die Frage, ob im Zeitalter des Anthropozäns überhaupt noch von MI als *Fähigkeit der Problemlösung* gesprochen werden kann.

Angelehnt an Wittgensteins Sprachspieltheorie ist es vielleicht unsinnig, zu fragen, was Intelligenz ist – wichtig ist, was man damit tut. Die Finalität natürlicher wie künstlicher Intelligenz muss untersucht werden und dabei von Intentionalität/Volition unterschieden werden, die allein dem wertenden und motiviertem Natürlichen System zukommen.

MI oder HI (human intelligence) wird hierbei als Teilbereich der Natürlichen Intelligenz (NI) aufgefasst und evolutionär als Resultat der Anpassungsnotwendigkeit und -fähigkeit auf sich ändernde Umweltbegebenheiten verstanden. Dieses Verständnis lässt die Weiterführung des Intelligenzbegriffs als Entwicklungsmöglichkeit neuer Fähig- und Fertigkeiten zu.

KI hingegen ist (im Anthropozän) als die von einer NI zielgerichtet entwickelten Funktionalität eines Werkzeugs, d. i. in ihrer Ausprägung messbar an dem Grad der Aufgabenerfüllung, aufzufassen. Somit ist jede KI immer der MI unterzuordnen. Ein anthropologischer Fehlschluss wäre, KI aufgrund der (programmierten) Finalität Intentionalität zu unterstellen.

⁹ Vgl. Eysenck (1970).

¹⁰ Vgl. Moor (2006).

Der 1990 eingeführte Begriff der Emotionalen Intelligenz sollte heute mit dem kognitiven (intellektuellen) Verständnis überein gebracht werden.

Emotionale Intelligenz ist, aufbauend auf Konzepten der Sozialen Intelligenz, die bis in die 1920er Jahre zurückreichen, ein von John D. Mayer (University of New Hampshire) und Peter Salovey (Yale University) im Jahr 1990 eingeführter Terminus. Er beschreibt die Fähigkeit, eigene und fremde Gefühle (korrekt) wahrzunehmen, zu verstehen und zu beeinflussen¹¹.

Das Konzept der emotionalen Intelligenz beruht auf der Theorie der multiplen Intelligenzen von Howard Gardner¹², deren Kerngedanke bereits von Edward Lee Thorndike und David Wechsler als „soziale Intelligenz“ bezeichnet wurde¹³.

Das Verständnis der Wichtigkeit der Sozialen Intelligenz ergänzt die Frage nach der Finalität von intelligentem Handeln um die Fragen der Intentionalität/Volition und damit Affekten wie auch Werten. So richtet sich intelligentes Verhalten implizit oder explizit nach Werten, nach ethischen Normen aus.

KI jedoch als per se wertfreie Intelligenzen, wenn sie auch *als implicit und/oder explicit ethical agents* sich augenscheinlich nach ethischen Richtlinien verhalten (können), sind nicht zu *full ethical agency* in der Lage und damit nicht zu affekt- und emotionsgetönten Agieren jenseits von Logik und Normen.

Ob dies tatsächlich ein Manko oder nicht vielmehr als positives Faktum zu verstehen ist, ist zu diskutieren. Die „gefühlte“ Freiheit, sich auch gegen (aus deontologischer Perspektive heraus) ethisch logisch-normativ-korrektes Verhalten entscheiden zu können, weil „einem der Sinn danach steht“, weil also „das Gefühl“ etwas anderes „diktiert“, sehe ich nicht als echte Wahlfreiheit, nicht als Freiheit, zwischen gleichwertigen Möglichkeiten zu wählen.

Die Realisierung der Möglichkeit, sich gegen ethische (Grund-)Überzeugungen (auch gegen die eigenen) zu stellen, führe ich vielmehr auf psychologische Phänomene, die ev. sogar Symptome einer Impulskontrollstörung imitieren, zurück. So ist die Aneignung eines bestimmten ethischen Grundmodells oft mit der Reduktion der kognitiven Dissonanz¹⁴ einhergehend, was jedoch noch nicht dazu führt, dass diese ethische Einstellung auch in kongruentes und spezifisches Verhalten um-

¹¹ Vgl. Mayer/Salovey (2004).

¹² Vgl. Gardener (1993).

¹³ Vgl. Thorndike (1920).

¹⁴ Vgl. Irle (2019).

gesetzt wird. Ein expliziter ethischer Agent, auch wenn er „nur“ auf Algorithmen basiert und nicht die Freiheit besitzt, sich gegen seine ethische Grundausrichtung zu verhalten, kann durchaus auch intentionsfrei ethisch agieren.

MI nutzt KI, um ihre eigene Effektivität und Effizienz zu steigern. Jede diskutierte und künstlich herbeigeführte Dichotomie zwischen MI und KI eröffnet einem Scheinproblem Raum. Auf einem Vergleich der Leistungsmöglichkeiten von menschlichem Intellekt und der Funktionalität von elektronischen Systemen zu beharren, kommt einem Vergleich eines Menschen mit einem Automobil gleich, der als Resultat die Aussage „Das Auto fährt schneller“ ergibt, und ist m.E. ähnlich unergiebig.

2. Verhält sich mensch im Anthropozän intelligent?

Schon in Bezug auf Natürliche (Menschliche) Intelligenz ist es herausfordernd, die bisher geltenden Begrifflichkeiten in Relation zum menschlichen Verhalten von heute zu setzen – Mensch 2.0 passt sich schon lange nicht mehr an die Umwelt als vielmehr die Umwelt an sich selbst an. Bei der Funktionalität Künstlicher *Intelligenzen* ist die Bestimmung eben derer als *intelligent* kontraintuitiv und kontraproduktiv. Hier löst das Verständnis der (vor-)gegebenen Verhaltensregeln von Künstlichen Systemen als *intelligent* im Sinne von autonom und willensbestimmt eine Flut von Projektionsfehlern aus, die bis zur Zuschreibung von Intentionalität, die sich in einer Vernichtungsabsicht der Menschheit niederschlagen soll, reichen.

Zunächst sollen an dieser Stelle dem Begriff der Intelligenz alle sich intelligent verhaltenden, im Sinne von: handlungs-zielorientierten Systeme unterstellt werden. Es wird die These vertreten, dass alle aktuell bestehenden intelligenten Systeme natürlicher Intelligenz untergeordnet sind, wobei MI mit anderen Animalischen Intelligenzen koexistiert und ihrerseits KI

- a. hervorgebracht hat und
- b. als Instrument/Werkzeug nutzt.

So ist im Rahmen von KI besser von deren Funktionalität statt von Intelligenz zu sprechen.

Bislang in der Psychologie diskutierte Konzepte menschlicher Intelligenz können nicht auf diese Funktionalität künstlicher Systeme angewendet werden. Auch MI

fordert im Anthropozän eine neue Betrachtungsweise und sollte um ein ethisch-pragmatisches Verständnis ergänzt werden, was meint, dass der seit den 1990er Jahren Verwendung findende Begriff der Emotionalen Intelligenz mit dem kognitiven (intellektuellen) Verständnis von Intelligenz überein gebracht werden muss.

Heute scheint es zweckmäßiger und wichtiger, sich nicht mehr mit Definitionen von Intelligenz aufzuhalten, sondern vielmehr Konzepte zu erarbeiten, angelehnt an Wittgensteins Sprachspieltheorie, sich zu fragen, was *mensch/maschine* mit dieser Intelligenz tut, wozu sie dient.

Die Gerichtetheit, die Finalität, natürlicher wie künstlicher Intelligenz muss untersucht werden und dabei von Intentionalität/Volition unterschieden werden.

MI als Teilbereich der Natürlichen Intelligenz zeigt sich m.E. evolutionär als Resultat der Anpassungsnotwendigkeit und -fähigkeit auf sich ändernde Umweltbegebenheiten (nach William Sterns Begriffsdefinition). Dieses Verständnis lässt die Weiterführung des Intelligenzbegriffs als Entwicklungsmöglichkeit neuer Fähigkeiten und Fertigkeiten zu.

KI (auch AI als Artificial Intelligence) hingegen wird von mir verstanden als die von einer NI zielgerichtet entwickelte Funktionalität eines Werkzeugs, d. i. in ihrer Ausprägung messbar an dem Grad der Aufgabenerfüllung. Somit ist KI der MI unterzuordnen. Ein anthropologischer Fehlschluss wäre, KI aufgrund der (programmierten) Finalität Intentionalität zu unterstellen.

In Bezug auf MI ist deren Finalität jedoch unbedingt mit Intentionalität/Volition verbunden und richtet sich damit implizit oder explizit nach ethischen Normen aus.

3. Kurzer historischer Abriss: Was ist „Intelligenz“?

Intelligenz ist als abstrakter Begriff umfasst neurobiologische, lerntheoretische und entwicklungsspezifische sowie testpsychologische Aspekte.

Das Verständnis von Intelligenz hat sich über die Jahrhunderte verändert, wobei bis heute keine einheitliche, allgemein anerkannte Definition von Intelligenz international Gültigkeit hätte¹⁵.

¹⁵ Zur Entwicklung des Intelligenzbegriffs und Geschichte des IQs vgl. u. a. Amelang/Bartussek (2001).

Dem am weitesten verbreiteten Verständnis von Intelligenz korrespondiert, auf Williams Sterns Begriffsbestimmung¹⁶ zurückgehend, dass intelligentes Verhalten bestmögliche Anpassung an sich ändernde Umweltbegebenheiten ist. 1912 beschreibt Stern (1871–1938) als Begründer des Intelligenzquotienten Intelligenz als die allgemeine Fähigkeit eines Individuums, sein Denken bewusst auf neue Forderungen einzustellen; sie ist allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens¹⁷.

Im Laufe der Jahrzehnte durchlief der Intelligenzbegriff zahlreiche Veränderungen¹⁸: Spearman forderte im 19. Jahrhundert eine differenziertere Betrachtung und plädierte für eine Unterscheidung von einem Generalfaktor der Intelligenz und zusätzlichen Spezialfaktoren. Wichtig ist zu beachten, dass zwischen Bildung und Intelligenz unterschieden werden muss.

Obwohl sich der menschliche Verstand in Hinblick auf rationale Fähigkeiten von dem aller anderen Lebewesen unterscheidet, entzog er sich lange Zeit jeder Vermessung und doch wurde über die Jahrtausende anhand bestimmter Fähigkeiten selektiert: So sollten Chinesen, die um 1100 v. Chr. für ihren Kaiser arbeiten wollen, sich in einem Talentwettbewerb beweisen, der Reiten, Bogenschießen, Musizieren, Schreiben und Rechnen umfasste. Im antiken Griechenland beschrieb ein unter dem Namen des Aristoteles rangierender Text die „Geistreichen“ anhand körperlicher Merkmale mit magerem Hals, schmalem Gesicht mit glanzvollen Augen, fleckenloser, weiß-rötlicher Haut¹⁹.

Der römische Gelehrte Marcus Tullius Cicero prägte 53 v. Chr. den Begriff der *intelligentia*: Intelligenz ist das Vermögen, das den Geist befähigt, die Wirklichkeit zu verstehen²⁰.

Im 19. Jahrhundert hielt die Vermessungslehre Einzug in die Intelligenzforschung²¹: Je größer die Gehirnmasse, desto größer seien Verstand und Charakter, vertrat irrtümlich der deutsche Arzt Franz Joseph Gall die phrenologische Auffassung.

¹⁶ Zum historischen Intelligenzverständnis vgl. Cornell University Library (2009).

¹⁷ Stern (1912).

¹⁸ Vgl. Wirtz (2020).

¹⁹ Vgl. Funke/Vaterrodt (2009), S. 13.

²⁰ <https://www.zeit.de/2013/13/Intelligenzforschung-Zeitleiste>, Zugriff: 17.10.2020.

²¹ Vgl. Oehler-Klein (1998).

Mit Einführung der Schulpflicht 1889 in Frankreich hatte sich die Heterogenität der Schülerschaft verstärkt. Insbesondere Kinder am unteren Ende des Begabungsspektrums, die vom Unterricht wenig profitierten, sollten daher zwecks gezielter Förderung ausgefiltert werden, um dann später wieder in die Regelschule integriert werden zu können. Bislang existierte noch kein einheitliches Verständnis davon, was Intelligenz überhaupt sei und ob bzw. dass dieses Konstrukt sich a) in menschlichen Fähigkeiten unterschiedlich niederschlägt und b) zu messen sei.

Der französische Psychologe Alfred Binet begann 1905 das Denkvermögen von Kindern anhand eines Fragenkatalogs zu testen²². In Abhängigkeit vom Alter der Kinder stellte er standardisierte Fragenkataloge zusammen, z. B.: Was reimt sich auf Pferd? Bilde einen Satz mit Paris, Fluss und Glück. Worin ähneln sich Fliege und Schmetterling? Dem Psychologen wird allerdings im Zuge der Testreihen klar, dass die letzte Frage nur soziologische Differenzen und keine Unterschiede in der Intelligenzleistung veranschaulicht: Nur Kinder wohlhabender Familien konnten diese Frage beantworten, da sie bei Landurlauben Schmetterlinge sehen; Arbeiterkinder aus der Stadt kennen sie nicht.

4. Intelligenzalter versus Lebensalter statt Reaktionsgeschwindigkeit und perzeptive Sensibilität

In England herrschte zu damaliger Zeit ein reduzierter Intelligenzbegriff vor: Während Francis Galton (ein Cousin Charles Darwins) in England die Ansicht vertrat, dass sich Intelligenz auf einfache mentale Operationen wie die Reaktionsschnelligkeit oder die Fähigkeit, zwischen Sinnesreizen wie Tonhöhen oder Helligkeit zu unterscheiden, herunterbrechen lässt, war Intelligenz für Binet abhängig von höheren kognitiven Funktionen: Aufmerksamkeit, Urteils- und Abstraktionsvermögen sind als Intelligenzleistungen ebenso ausschlaggebend für den Schulerfolg²³.

Binet legte eine Definition zu Grunde, die Intelligenz als die Fähigkeit zur Bewältigung alltäglicher Probleme erachtet. Die Aufgaben wurden nun nach Schwierigkeit hierarchisch geordnet und gestaffelt. Diejenigen, die von jeweils 70 % der Kinder eines Altersjahrgangs gelöst werden konnten, wurden zu „Altersreihen“ zu-

²² Vgl. Lamberti (2005).

²³ Vgl. ebd.

sammengefasst. Kann ein Kind alle Aufgaben seiner Altersstufe lösen, entspricht sein Intelligenzalter seinem Lebensalter; löst das Kind weitere Aufgaben, liegt sein Intelligenzalter entsprechend darüber. Als problematisch sah es Binet selbst allerdings an, den kognitiven Entwicklungsstand eines Kindes mit einer Zahl zu beschreiben, da dieselbe Zahl gelöster Aufgaben sehr unterschiedliche Muster an Fähigkeiten beinhalten kann.

Hinzu kam, dass gleiche Differenzen zwischen Intelligenz- und Lebensalter bei Kindern unterschiedlichen Alters nicht dasselbe bedeuten. Ein Fünfjähriger mit einem Intelligenzalter von sieben Jahren ist seinen Altersgenossen in seiner Entwicklung deutlich weiter voraus als ein Achtjähriger auf dem Niveau eines Zehnjährigen.

William Stern relativierte nun die Intelligenz am Lebensalter, statt die Differenz als Maß für die Entwicklungsbeschleunigung oder -verzögerung zu nehmen. Dieser „Intelligenzquotient“ wurde dann zur leichteren Handhabbarkeit noch einmal mit 100 multipliziert – stimmen Intelligenz- und Lebensalter überein, ergibt sich somit ein IQ von 100 als Durchschnittswert. Diese Art der Messung erklärt auch, wie die teilweise astronomischen Intelligenzwerte höchstbegabter Kinder zustande kommen. Ein IQ von 200 beispielsweise ergibt sich nach Sterns Formel dann, wenn das Intelligenzalter doppelt so hoch ist wie das Lebensalter. Die amerikanische Weiterentwicklung von Binets Test durch Lewis Terman (der später dann seine berühmte Hochbegabten-Längsschnittstudie initiierte), der Stanford-Binet-Test, wird mangels Alternativen teilweise noch heute in der Höchstbegabtenforschung angewandt.

Wie aber kann man Intelligenz messen, wenn sich Lebensalter und zu lösende Aufgaben nicht mehr so eindeutig zuordnen lassen? Entwicklungspsychologische Befunde und schulische Lehrpläne erlauben dies für die Kindheit; mit zunehmendem Alter werden die Fähigkeitsunterschiede zwischen Menschen immer differenzierter, denn auch unterschiedliche Berufe implizieren unterschiedliches Wissen. Intelligenz ist jedoch nicht mit angesammeltem Wissen gleichzusetzen.

Aus diesem Grund schlug David Wechsler 1939 vor, keine konkreten Altersreihen mehr zum Maßstab zu nehmen, sondern zu ermitteln, inwieweit der Testwert einer Person von den Werten der Gesamtgruppe abweicht²⁴. Wie viele

²⁴ Vgl. ebd.

Merkmale ist auch die Intelligenz normalverteilt – die Glockenkurve ist symmetrisch, und mehr Personen erreichen Werte im mittleren als im extremen Bereich. Um einen Test so zu eichen, dass eine genaue Messung möglich ist, ist eine umfangreiche Normierungsstichprobe erforderlich. Die in Deutschland gängige Skalierung behielt den von Stern ins Leben gerufenen Mittelwert von 100 bei, der die Intelligenzverteilung in zwei Hälften teilt; mit einer Standardabweichung von 15 lassen sich Unterschiede hinreichend akkurat beschreiben. Das gängige 130er Kriterium für Hochbegabung ergibt sich somit rein statistisch: Zwei Standardabweichungen über dem Mittelwert gelten als hinreichend überdurchschnittlich, so dass man von Hochbegabung sprechen kann. Auch wenn der heutige IQ streng genommen also gar kein Quotient mehr ist, ist der Name, den Stern ihm vor 100 Jahren gab, geblieben. Stern indes unterschätzte den Wert seiner methodischen Innovation doch ein wenig: In seiner Autobiographie findet der IQ noch nicht einmal Erwähnung.

Und doch ist bis dato unklar, was Intelligenz über kognitive (Wissens-)Leistungen hinaus ist. Sterns Idee, die bis heute Gültigkeit hat, umfasst das Verständnis des IQ als „bestmögliche Anpassungsfähigkeit an sich ändernde Umweltgegebenheiten“.

5. Der g-Faktor der Intelligenz

Charles Spearman (1863–1945) entwickelte mit Hilfe des Generalfaktorenmodells der Faktorenanalyse seine Zwei-Faktoren-Theorie. Beim Generalfaktormodell wird aus den vorliegenden Testdaten ein einziger übergeordneter Faktor extrahiert. Auf diese Weise extrahierte Spearman einen Faktor für allgemeine Intelligenz, den *g-factor* (*g* von *general*). Diesen Faktor beschrieb er als generellen, alle unterschiedlichen Leistungsbereiche beeinflussenden Intelligenzfaktor. Die Ausprägung dieser allgemeinen Intelligenz bestimme u. a. Verarbeitungsgeschwindigkeit, geistige Kapazität, intellektuelle Leistung – also insbesondere, ob eine Person eher schnell oder langsam denkend ist.

Neben diesem *g-factor* extrahierte Spearman noch eine Reihe von untergeordneten, spezifischen Intelligenzfaktoren, welche er als *s-factors* (*s* von *specific*) bezeichnete. Diese sind dem *g-factor* hierarchisch untergeordnet und bereichsspezifische, voneinander unabhängige Faktoren. Sie bestimmen (jedoch

wesentlich vom *g-factor* beeinflusst) die Leistung einer Person in bestimmten Bereichen – z. B. bei mathematischen Aufgaben, bei verbalen oder räumlichen Problemstellungen.

6. Kristalline und fluide Intelligenz

Seine Vorstellungen wurden von Cattell (1905–1998) aufgegriffen und präzisiert. Cattell unterschied zwei Faktoren zweiter Ordnung, die er als fluide und kristalline Intelligenz bezeichnete. Die fluide Intelligenz repräsentiert die gehirnphysiologische Effizienz, die sich beispielsweise in der Verarbeitungsgeschwindigkeit ausdrückt. Dagegen ist die kristalline Intelligenz als Niederschlag individueller Erfahrungen zu sehen, die etwa dem verbalen Verstehen oder der routinisierten Durchführung effektiver Problemlösestrategien zugrunde liegen. Während die Entwicklung der kristallinen Intelligenz somit stark sozialisationsabhängig ist, nahm Cattell fluide Intelligenz als genetisch fixiert an.

Thurstone (1887–1955) nahm die Position ein, dass Intelligenzleistungen das Zusammenspiel von sieben unabhängigen Primärfaktoren widerspiegeln, darunter verbale Kompetenzen, Merk- und Rechenfähigkeit sowie räumliches Vorstellungsvermögen. Auch den Begriff der sozialen Intelligenz führte er ein.

Mit emotionaler Intelligenz bezeichnen John D. Mayer und Peter Salovey die Fähigkeit, eigene und fremde Gefühle wahrzunehmen, zu verstehen und zu beeinflussen. Als Basisgedanken der emotionalen Intelligenz entwickelten schon Anfang des letzten Jahrhunderts die Testpsychologen Edward Lee Thorndike und David Wechsler den Begriff der sozialen Intelligenz.

Guilford ging dagegen von einer Vielzahl weiterer Einzelfaktoren aus (insgesamt 150), zu denen er aufgrund theoretischer Vorüberlegungen gelangte. In seinem Würfelmodell entsprechen die drei räumlichen Dimensionen des Würfels (1) intellektuellen Operationen (z. B. Erkenntnis, Gedächtnis, Bewerten), (2) deren Inhalte (z. B. figural, symbolisch) und (3) den Produkten intelligenter Verhaltens (z. B. System, Transformation). Während Guilford seinem Modell v. a. eine heuristische Funktion zuerkannte und durch seine Unterscheidung von konvergenter (Intelligenz) und divergenter Denkproduktion (Kreativität) bekannt wurde, rekurrieren sogenannte Intelligenzstrukturtests oder differentielle Fähigkeitstests häufig auf Thurstones Primärfaktorenmodell.

7. Entwicklung des IQ in der Lebensspanne

In der 2. Grundschulklasse sind 80 % der Intelligenz bereits entwickelt. Ab dem 26. Lebensjahr nimmt die fluide Intelligenz ab, die kristalline bleibt konstant bzw. kann weiter (durch Bildung, Erfahrung, Training) gesteigert werden. Hans Eysenck erforschte die Unterschiede in der Persönlichkeit von Menschen, wobei er schon in seiner ersten Buchpublikation anno 1947, *Dimensions of Personality*, das Konzept von Introvertiertheit, Extravertiertheit und Neurotizismus entwickelte²⁵. Er ging davon aus, dass diese Merkmale genetisch bedingt und in der Physiologie des Menschen verankert sind, wobei er deren Sitz im Gehirn und im zentralen Nervensystem vermutete. Durch unterschiedliche Stufen der Erregung des Gehirns werden entweder die Extravertiertheit oder die Introvertiertheit ausgebildet, wird emotional auf Ereignisse reagiert, tritt die neurotische Dimension in den Vordergrund, die ihre Basis im Nervensystem hat.

Turhan Canli von der Stony Brook University, New York konnte schon in den 2000ern zeigen²⁶, dass die Amygdala, Teil des Limbischen Systems und damit auch emotionsregulierend, bei extravertierten Probanden linksseitig vergrößert ist. Umgekehrt geht die Tendenz zu negativen Gefühlen (entspricht dem Konzept des Neurotizismus) rechtsseitig mit einem verringertem Amygdalavolumen einher.

Persönlichkeit ist damit die mehr oder weniger stabile und dauerhafte Organisation des Charakters, Temperaments, Intellekts und Körperbaus eines Menschen, die seine einzigartige Anpassung an die Umwelt bestimmt. Der Charakter eines Menschen bezeichnet das mehr oder weniger stabile und dauerhafte System seines konativen Verhaltens (des Willens); sein Temperament das mehr oder weniger stabile und dauerhafte System seines affektiven Verhaltens (der Emotion oder des Gefühls); sein Intellekt das mehr oder weniger stabile und dauerhafte System seines kognitiven Verhaltens (der Intelligenz); sein Körperbau das mehr oder weniger stabile System seiner physischen Gestalt und neuroendokrinen (hormonalen) Ausstattung²⁷.

Die Darstellung der im Laufe der Jahrhunderte sich verändernden Konzeptionen von Intelligenz soll verdeutlichen, dass Intelligenz relativ ist, ein Konstrukt, kein naturgegebenes Merkmal.

²⁵ Vgl. Eysenck (1997).

²⁶ Vgl. Canli u. a. (2002).

²⁷ Eysenck (1970), S. 2.

8. Der EQ: der größte Unterschied zur KI

Daniel Goleman definiert den Begriff der emotionalen Intelligenz²⁸ in Anlehnung an Salovey und Gardner²⁹ durch die folgenden Fähigkeiten:

Die eigenen Emotionen kennen

Die eigenen Emotionen erkennen und akzeptieren, während sie auftreten. Diese Fähigkeit ist entscheidend für das Verstehen des eigenen Verhaltens und der eigenen Antriebe. (Hintergrund: Viele Menschen fühlen sich gegenüber ihren Gefühlen ausgeliefert, lehnen sie ab und bekämpfen oder vermeiden sie – statt sich der Tatsache bewusst zu sein, dass man Emotionen aktiv steuern kann.)

Emotionen beeinflussen

Gefühle so handhaben, dass sie der Situation angemessen sind (statt zu dramatisieren oder zu verharmlosen). Dazu gehört die Fähigkeit, sich selbst zu beruhigen und Gefühle der Angst, Gereiztheit, Enttäuschung oder Kränkung abzuschwächen und positive Gefühle zu verstärken. Dies hilft bei der Überwindung von Rückschlägen oder belastenden Situationen.

Emotionen in die Tat umsetzen

Emotionen beeinflussen derart, dass sie bei der Erreichung von Zielen helfen. Dies ist der Kern der Selbstmotivation und fördert die Kreativität sowie die Häufigkeit von Erfolgserlebnissen. Dazu gehört auch, dass jemand in der Lage ist, kurzfristige (emotionale) Vorteile und Verlockungen hinauszuschieben (Belohnungsaufschub) und impulsive Reaktionen zu unterdrücken. Diese längerfristige Perspektive ist die Grundlage jeglichen Erfolges. In späteren Arbeiten ordnete Goleman diesen antriebs- und motivationsbezogenen Aspekt der oben genannten Fähigkeit Emotionen zu beeinflussen zu.

²⁸ Vgl. Goleman (1997), S. 65ff.

²⁹ Vgl. Salovey (2004).

9. Empathie und Intelligenz

Als Giacomo Rizzolatti's Entdeckung der Spiegelneurone in den 1990er Jahren die Psychologie und Biologie aufschreckte, hatte ich das Glück, bei seinen ersten Vortragsreisen im Publikum sitzen zu dürfen: Auch, wenn Rizzolatti damals selbst noch vor voreiligen Schlüssen und unbedachten Übertragungen seiner (aus Affenexperimenten gewonnen) Erkenntnisse warnte, gilt heute doch als hinreichend belegt, dass eben jene Spiegelneurone die biologische Basis für Mitgefühl und Empathie abbilden³⁰.

Empathie ist die Grundlage von Menschenkenntnis (und -verständnis) und das Fundament zwischenmenschlicher Beziehungen. Ein Mensch, der erkennt, was andere fühlen, kann Signale im Verhalten anderer erkennen, deuten und interpretieren, um Motivation und Intention des Anderen abzuleiten. Empathie selbst ist als eine wertneutrale Fähigkeit zu verstehen – sie kann individuell positiv oder auch negativ empfundene Wirkung haben.

Die Fähigkeit der erfolgreichen Gestaltung von Beziehungen besteht essentiell im Umgang mit den Gefühlen anderer Menschen. Empathie ist Grundlage für eine möglichst reibungslose Zusammenarbeit in nahezu allen beruflichen Umfeldern. Gleichzeitig gilt sie als Voraussetzung für Beliebtheit, Wertschätzung und Integration in eine Gemeinschaft (*Soziabilität*), andererseits aber auch für *leadership ability*, der Führungsqualität.

Zur Operationalisierung und Messung der emotionalen Intelligenz in einem Test haben Salovey und Mayer dieses Konzept in vier Bereiche gegliedert:

- Wahrnehmung von Emotionen
- Nutzung von Emotionen
- Verstehen von Emotionen
- Beeinflussung von Emotionen

Wahrnehmung von Emotionen:

- Emotionen in Gesichtern identifizieren
- Emotionen in Landschaften und Designs identifizieren

³⁰ Vgl. Rizzolatti (2008).

Nutzung von Emotionen zur Unterstützung des Denkens:

- emotionale Empfindungen mit anderen taktilen oder sensorischen Stimuli vergleichen
- Emotionen identifizieren, die bestimmte Denkaufgaben bestmöglich unterstützen

Verstehen von Emotionen:

- Wissen, unter welchen Umständen emotionale Zustände wechseln und wie ein emotionaler Zustand in einen anderen übergeht
- Mehrere Emotionen in komplexeren affektiven Zuständen identifizieren

Umgang mit Emotionen:

- Maßnahmen zur Veränderung des eigenen emotionalen Zustandes in hypothetischen Szenarien vorschlagen
- Maßnahmen zur Beeinflussung des emotionalen Zustands anderer Personen zur Zielerreichung vorschlagen

Der EQ ist bei höheren Säugetieren, die wie Menschen in Herden/Rudeln organisiert leben, ausschlaggebend für die soziale Integration und damit ausschlaggebend für die individuelle Existenzsicherung. Welchen Sinn hat es somit, Intelligenz nur als rein kognitive, intellektuelle Leistungsfähigkeit zu verstehen? Zudem erweist es sich seit Mitte der 1990er Jahre, dass eben diese Intelligenzleistungen, zumindest in Teilbereichen, transgenerational wieder abnehmen:

10. Intelligenz im Anthropozän: Der Anti-Flynn-Effekt

Ein Forscherteam vom King's College London legte 2017 eine neue Studie vor, nach der seit 1972 zunehmend schlechter in einem Teilaspekt von IQ-Tests abgeschnitten wird³¹: Bei Aufgaben, die das Arbeitsgedächtnis testen, sinkt die Leistungsfähigkeit kontinuierlich. Gleichzeitig haben sich Resultate für das Kurzzeitgedächtnis gesteigert. Wongupparaj u. a. analysierten als Versuch, die Ergebnisse der zum Teil unterschiedlichen IQ-Tests vergleichbar zu machen, insgesamt

³¹ Wongupparaj u. a. (2017), S. 71–80.

1754 IQ-Tests, mit Fokus auf die Ergebnisse für die Teilbereiche Arbeits- und Kurzzeitgedächtnis.

Das Arbeitsgedächtnis bezeichnet eine komplexere kognitive Fähigkeit (inkludiert das Kurzzeitgedächtnis, kann drüber hinausgehend aber dessen Inhalte manipulieren). Informationen werden im Arbeitsgedächtnis nicht nur repräsentiert, sondern auch manipuliert. Als Beispiel: Wenn im Supermarkt gedanklich die Einkaufsliste abgerufen wird und mit spontanen Plänen fürs Abendessen verglichen wird, ziehen wir verschiedene Informationen aus dem Gedächtnis und der Umwelt zusammen, wägen ab und treffen eine Entscheidung.

Die Beobachtung, dass die Ergebnisse fürs Kurzzeitgedächtnis über die Jahrzehnte besser wurden, entspricht dem Flynn-Effekt³². Danach schneiden Menschen in den Industrienationen über die Generationen bei (nicht nachgeachteten) IQ-Tests zunehmend besser ab. Bis in die 1990er Jahre wurden damit Menschen aus Industrienationen offenkundig „intelligenter“. Oft wird dies auf die Verbesserung von Bildung, Ernährung und Gesundheitsversorgung zurückgeführt, könnte aber auch einen Lerneffekt im Umgang mit IQ-Test-Aufgaben darstellen.

Die Frage, ob sich der Flynn-Effekt seit den 1990ern umkehrt und die Menschen an Intelligenz einbüßen, also „dümmer“ werden, ist so nicht zu beantworten. Die Analyse von Wongupparaj et al. könnte in Teilen dafür sprechen, dass bestimmte kognitive Fähigkeiten durch eine Minderbeanspruchung oder teils sogar Nichtnutzung (durch Eskalation an technische Hilfsmittel) zumindest untrainiert sind.

Allerdings stellten die Forscher in der Studie auch fest, dass der Anteil von Über-60-Jährigen, die an IQ-Tests teilnehmen, über die Zeit größer wurde. Das Arbeitsgedächtnis gehört zu den kognitiven Funktionen, die im Alter nachlassen. Dieser Stichprobeneffekt könnte für die AG-Ergebnisse mitverantwortlich sein – und auch andere Studien erklären, nach denen der IQ in den Industrienationen sinkt.

Manfred Spitzer nennt u. a. den vermehrten Medienkonsum als mitverursachend für die Stagnation und spricht hier sogar von der digitalen Demenz³³.

Die These, dass die vermehrte Migration aus Ländern mit durchschnittlich niedrigerem IQ für den Anti-Flynn-Effekt verantwortlich zeichnet, widerlegen J. Pietchnig u. a. von der Universität Wien. Sie weisen nach, dass binnen ein bis zwei Generationen der IQ sich an den Durchschnittswert des Gastlandes anpasst – in

³² Flynn (2007).

³³ Vgl. Spitzer (2014).

beide Richtungen, nach oben wie nach unten³⁴. Dies wiederum entspricht dem traditionellen Verständnis von Intelligenz als Anpassungsfähigkeit an sich verändernde Umweltbedingungen: Sogar die Intelligenzleistung selbst passt sich ihre Umgebung an.

Wenn Eysencks Verständnis der Intelligenz als Persönlichkeitsvariable mitberücksichtigt wird, könnte der Anti-Flynn-Effekt auch auf die zunehmende Einschränkung der realen Interaktion und sprachlichen Kommunikation zugute der digitalen Kommunikation und der „Bildsprache“ durch Emoticons rückzuführen sein.

Dies bedeutet auf der anderen Seite, dass nicht nur die akademische Intelligenz, sondern auch und gerade die Emotionale Intelligenz in ihrer weiteren transgenerationalen Entwicklung untersucht werden muss. Hat die vermehrte Nutzung von KI als z. B. Navigationsgeräte, als Übersetzungs- und Rechenhilfe, als Kommunikationsassistenten etc. einen Einfluss auf den EQ der Nutzer und wie wirkt(e) sich der vermehrte Medienkonsum (auch TV, Computer etc.) der Bezugspersonen von Neugeborenen in der sensiblen Entwicklungsphase der Entstehung des Empathievermögens seit Mitte der 1990er Jahre aus?

Hier lohnt es sich, Langzeitstudien, die in Folge von M. Main und J. Solomon³⁵ stehen, die ihrerseits auf den Attachment-Studien von M. Ainsworth et al. aufbaute, zu initiieren: Wie hat sich das als desorganisiert/desorientiert eingestufte Bindungsverhalten der Kinder weiter entwickelt, und in welchem Zusammenhang ist Attachment mit der Emotionalen Intelligenz (auch und gerade als Fähigkeit, eigene und fremde Emotionen zu erkennen und zu regulieren) zu sehen? Ist das als „still face“ beschriebene Gesicht der Bezugspersonen beim Medienkonsum als ein Faktor emotionaler Vernachlässigung aufzuführen?

Ist die eingangs von mir vermutete Zunahme affektiver Störungen und Impulskontrollstörungen evtl. auch auf die fehlende organismische (und damit affekt- und emotionsbeeinflussende) Ko-Regulierung durch (medien-absorbierter) Bezugspersonen im sensiblen Entwicklungsalter zurückzuführen?

Wenn Eysencks Verständnis der Intelligenz als Persönlichkeitsvariable mit berücksichtigt wird, könnte der Anti-Flynn-Effekt auch auf die zunehmende Einschränkung der realen Interaktion und sprachlichen Kommunikation zugute der

³⁴ Pietschnig (2018).

³⁵ Main u. a. (1986), S. 95–124.

digitalen Kommunikation und der „Bildsprache“ durch Emoticons rückzuführen sein.

11. Fazit: Der Mensch als dümmstes und unglücklichstes Wesen

MI ist als Teil der NI zu sehen, die Werkzeuge und Instrumente hervorbringt, mittels derer sie ihre Effizienz und Effektivität steigern kann. Aus dieser Perspektive heraus und Pietschnigs Studien zur Migrationsforschung berücksichtigend, die ergaben, dass Intelligenz sich transgenerational an den Durchschnitt adaptiert, kann durchaus behauptet werden, dass mensch im Anthropozän gerade durch die Delegation und Eskalation bestimmter Aufgaben an technische und technologische Hilfsmittel weiter intelligentes Verhalten zeigt.

MI ist variabel, anpassungsfähig – sowohl transgenerational als auch individuell und umweltbedingungsabhängig.

Intelligenz als Möglichkeit, individuelle Bestleistungen zu bewirken, ist dabei abhängig von Motivation, Affekt und Werten.

KI ist als vom Menschen geschaffenes Werkzeug, dessen Leistungsfähigkeit als Grad von *Funktionalität* zu messen ist, zu sehen. William Sterns Verständnis von Intelligenz scheint heute obsolet, da der Mensch seine Umwelt an sich anpasst, teils sogar natürliche Ressourcen irreversibel zerstört.

Intelligenz als Persönlichkeitsmerkmal zu betrachten scheint heute zutreffender. Persönlichkeit ihrerseits zeigt sich jedoch in Interaktion, außerhalb von Beziehungen spielen Persönlichkeitsmerkmale, wenn überhaupt, nur eine untergeordnete Rolle. Nur relativ zu den Merkmalen anderer Individuen sind individuelle Persönlichkeitseigenschaften interessant. Als Persönlichkeitsvariable ist MI zweckgebunden und zielorientiert.

KI kennzeichnet sich weder durch Persönlichkeit noch durch Intentionalität. Der EQ wiederum steht in Zusammenhang mit dem Verständnis vom vollständig ethisch handelnden Agens, eine Fähigkeit, die KI aberkannt werden muss³⁶. KI ist als implizit ethischer Agens in der Lage, Programmanweisungen, die ethische Fragestellungen implizieren, zu folgen. Als explizit ethischer Agens ist sie heute teils in der Lage, nach ethischen Richtlinien eigenständig zu entscheiden:

³⁶ Moor (2006).

A machine that is able to calculate the best action in ethical dilemmas using an ethical principle, as opposed to having been programmed to behave ethically, where the programmer is following an ethical principle³⁷.

Ein explizit ethischer Agent nutzt deontische Logik für Entscheidungen bzw. Aussagen über Erlaubnis und Verpflichtung, epistemische Logik für Aussagen über Glauben und Wissen, sowie Handlungslogik für Aussagen über Handlungen.

Die deontische Logik versieht die Maschine mit ethischen Imperativen, an denen Situationen und/oder Handlung geprüft werden können. Mittels epistemischer Logik kann die Validität des vorhandenen Wissens beurteilt werden und berücksichtigt werden. Die Handlungslogik ermöglicht die ethische Bewertung von Handlungen und ist damit entscheidungsbestimmend.

Der Mensch als vollständig ethischer Agent bettet seine logischen Ableitungen dabei immer in seinen persönlichen Kontext ein, weiß um seine Wahlmöglichkeiten und -freiheiten und agiert gemäß seinen Vorlieben etc. – hier zeigt sich der Mensch erneut affektbestimmt³⁸. Metahuman existiert keine Intentionalität, kein Wille – KI will nichts, hat also die Zielverfolgung nicht an affektbedingte Motivation gekoppelt.

Wille (als strategische Motivation über Instinkte, Reflexe hinaus) ist dabei auch an Werte gebunden. Ein entwickelter Wertecodex, der bewusst reflektiert wird, ist die Grundlage für jedes deontologisch Verhalten. Werte jedoch entstehen aus natürlichen Grundlagen, sind Ergebnis unseres sozialen Menschseins, und sie spiegeln die prinzipiellen Funktionalitäten unseres Sozialorgans Gehirn wider: Wir Menschen sind, wie alle bewusstseinsfähigen Lebewesen überhaupt, zur Empathie fähig und von ihr bestimmt. Wenn KI nichts WILL, sie wird auch nicht überleben wollen, da sie (materielle/biologische) Existenz und Nichtexistenz nicht als gut oder schlecht (be-)wertet. (Kann ein „Gott“ deshalb auch Leiden und Elend der materiell-biologischen Welt ertragen, weil es für ihn keine Bewertung und damit kein Leid und Elend an sich gibt?)

Der Intelligenzbegriff aus psychologischer Sicht ist, zusammenfassend, relativ und als Konstrukt zu verstehen. Intelligenz umfasst dabei sowohl akademisch-intellektuelle Fähigkeiten, die Reaktionsfähigkeit und damit Funktionalität des Gehirns wie auch soziale Kompetenzen, inkludiert also IQ wie EQ.

³⁷ Anderson (2007).

³⁸ Moor (2006).

Im Anthropozän beraubt sich *mensch* seiner Lebensgrundlage: einer ausbalancierten Natur, die sich selbst, in kontinuierlicher Veränderung und Anpassung, stabilisiert.

Die Psychologie verstand unter Intelligenz seit Anfang 1900 nach William Stern bestmögliche Anpassung an sich ändernde Umweltbegebenheiten, Fähigkeit zur Anpassung an unbekannte Situationen bzw. zur Lösung neuer Probleme: Der Mensch jedoch steht nun vor der Aufgabe, selbst gemachte Probleme zu lösen bzw. die Umwelt erneut so zu verändern, dass er überhaupt wieder die Chance hat, sich an die Umwelt anzupassen. Der Vergleich des Zauberlehrlings mag hier heranzuziehen sein, der die Geister, die er rief, nicht mehr kontrollieren kann. Nun will und hofft mensch, mit Hilfe von KI diese hausgemachten Probleme zu lösen und übersieht hierbei seine eigene Unfähigkeit, eine Intelligenz intentional und kontrolliert zu schaffen, die der eigenen insofern überlegen sein könnte (also Probleme zu lösen vermochte, die vom Menschen selbst nicht zu lösen sind), dass sie nach anderen Prinzipien gerichtet wäre.

Bleibe einer gleich wie gearteten Intelligenz, die in der Fähigkeit, sich die Umwelt, dem Menschen ebenbürtig, Untertan zu machen, überhaupt eine andere Entscheidung, als den Parasiten Mensch vom Planeten zu eliminieren? Wäre es nicht Zeichen einer NI, das, was die Umwelt zerstört (das Verhalten des Menschen) auszuschalten? MI wird hierbei als Teilbereich der NI und evolutionär entstanden als Resultat der Anpassungsnotwendigkeit und -fähigkeit auf sich ändernde Umweltbegebenheiten aufgefasst. Dieses Verständnis lässt die Weiterführung des Intelligenzbegriffs als Entwicklungsmöglichkeit durchaus neuer Fähig- und Fertigkeiten zu.

KI hingegen wird von mir betrachtet als die von einer NI zielgerichtet entwickelte Funktionalität eines Werkzeugs, d. i. in ihrer Ausprägung messbar an dem Grad der Aufgabenerfüllung. Somit ist KI MI unterzuordnen. Ein anthropologischer Fehlschluss bzw. ein Projektionsfehler wäre, KI aufgrund der (programmierten) Finalität Intentionalität zu unterstellen. In Bezug auf MI ist deren Finalität jedoch unbedingt mit Intentionalität/Volition verbunden und richtet sich damit implizit oder explizit nach ethischen Normen aus. MI muss unter dem finalitätsgerichteten Aspekt von Intentionalität und Volition untersucht werden: Der Mensch nutzt seine (und die von ihm entwickelte künstliche) Intelligenz, um zu überleben bzw. um das Leben so angenehm wie möglich zu gestalten – KI hingegen beabsichtigt und will nicht sowie nichts. Hier wird die ethische Implikation deutlich.

Zeigt sich aus diesem Verständnis heraus der Mensch im Anthropozän als intelligent? Der Mensch passt nicht mehr sich und seine Handlungen an die Umwelt an, sondern er verändert die Umwelt derart, dass er selbst unter gegebenen Bedingungen keine neuen Fähigkeiten mehr entwickeln und bereits erworbene nicht mehr nutzen muss (z. B. Lesen, Schreiben, Raumorientierung, soziale Fähigkeiten etc.) Auf der anderen Seite beraubt er sich (aus hedonistischen Motiven heraus?) durch die fortlaufende Beeinträchtigung und Beeinflussung natürlicher Prozesse seiner Lebensgrundlage Natur.

Aus dieser Perspektive heraus ist das Verständnis von Intelligenz als Fähigkeit zur Lösung neuer Probleme nicht mehr gültig. Anstatt sich nun auf die Lösung der hausgemachten Problematiken zu konzentrieren (z. B. könnte KI auch als APP dahin gehend genutzt werden, dass der ökologische Fußabdruck von Waren als ad hoc Kaufentscheidungshilfe berechnet werden kann) müht sich der Mensch an Spekulationen über die Eventualität zukünftiger Bedrohung durch die von ihm geschaffene KI ab. Er befürchtet eine apokalyptische Überwältigung durch autonome KI, die dem Kontrollverlust im Zauberlehrling gleicht. Hier zeigt sich ein anthropologischer Gedankenfehler: KI ist in Intentionalität und ihrer Finalität eben nicht mit MI zu vergleichen und ist intentions- wie volitionsfrei. Ebenso wenig bewertet KI von sich aus (Über-)Leben als gut und/oder Untergang/Tod als schlecht. Von KI ist mithin weder Unterstützung noch eine Gefährdung hinsichtlich des Fortbestands der Spezies Mensch zu erwarten.

Unter Berücksichtigung alles im Vorfeld Dargestellten plädiere ich für einen neuen, ethisch-pragmatischen Intelligenzbegriff. Hier sollte und muss auch Emotionale Intelligenz mit berücksichtigt werden. Bei künstlichen Systemen als Werkzeug/Instrument ist auf den Begriff der Intelligenz zugunsten des Begriffs der Funktionalität zu verzichten. In der Definition von MI muss deren Intentionalität und damit moralische Ausrichtung im Sinne des (Überlebens-)Schutzes einer jeder Form von Intelligenz (auch natürliche und animalische Intelligenz) stärker gewichtet werden. Ich gerate immer mehr zu der Einschätzung, dass der Mensch nicht nur das wohl unglücklichste, sondern auch das dümme (im Sinne von nicht intelligent handelndem) Lebewesen dieses Planeten ist und sich selbst zu eben jenem macht, seit er die natürliche durch die von ihm gestaltete künstliche Evolution zu überwinden sucht.

Literatur

- Amelang, Manfred/Bartussek, Dieter (⁵2001): *Differentielle Psychologie und Persönlichkeitsforschung*, Stuttgart
- Anderson, Michael (2007): Machine Ethics: Creating an Ethical Intelligent Agent, in: *AI Magazine* 28/4, S. 24
- [Pseudo-]Aristoteles (2006): *Physiognomik*, Waldshut
- Bartneck, Christoph u. a. (2019): *Ethik in KI und Robotik*, München
- Canli, Turhan u. a. (2002): Amygdala response to happy faces as a function of extraversion, in: *Science* 296/5576, S. 2191, PubMed PMID: 12077407
- Cattell, Raymond Bernard (1937): *The fight for our national intelligence*, London
- Cornell University Library (2009), *The Psychological Methods of Testing Intelligence* [1914], Cornell
- Dogs, Christian Peter (2017): *Gefühle sind keine Krankheit*, Berlin
- Eysenck, Hans Jürgen (1970): *The structure of human personality*, London
- (1997): *Dimensions of personality*, London
- Flynn, James R. (1987): Massive IQ gains in 14 nations: What IQ tests really measure, in: *Psychological Bulletin* 101, S. 171–191
- (2007): *What is intelligence? Beyond the Flynn effect*, Cambridge, UK
- Funke, Joachim/Vaterrodt, Bianca (2009): *Was ist Intelligenz?*, München
- Gardener, Howard (1993): *Multiple intelligences: the theory in practice*, New York
- (2013): *Intelligenzen: Die Vielfalt des menschlichen Geistes*, Stuttgart
- Goleman, Daniel (1997): *Emotionale Intelligenz*, München
- Irle, Martin, Hg. (2019): *Theorie der Kognitiven Dissonanz*, Göttingen
- Kaplan, Andreas M./Haenlein, Michael (2019): Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence, in: *Business Horizons* 62/1, S. 15–25
- Lamberti, Georg, Hg. (2005): *Intelligenz auf dem Prüfstand. 100 Jahre Psychometrie*, Göttingen
- Main, Mary u. a. (1986): Discovery of an insecure-disorganized/disoriented attachment pattern, in: Thomas Berry Brazelton/Michael Yogman, Hg., *Affective development in infancy*, Westport CT, S. 95–124
- Mangal, S K (2009): *General Psychology*, New York
- Mayer, John D./Salovey, Peter (2004): *Emotional Intelligence: Key Readings on the Mayer and Salovey Model*, Naples
- Moor, James H. (1995): Is Ethics Computable?, in: *Metaphilosophy* 26/1–2, S. 1–21
- (2006): The Nature, Importance, and Difficulty of Machine Ethics, in: *IEEE Intelligent Systems*, S. 18–21

- Moravec, Hans (1999): Rise of the Robots, *Scientific American*, December 1999, S. 124–135
- Oehler-Klein, Sigrid (1998): *Die Schädellehre Franz Joseph Galls in Literatur und Kritik des 19. Jahrhunderts*, München
- Pietschnig, Jakob u. a. (2018): Is the Flynn effect related to migration? Meta-analytic evidence for correlates of stagnation and reversal of generational IQ test score changes, in: *Politische Psychologie* 2, S. 267–283
- Pietschnig, Jakob/Gittler, Georg (2015): A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT-based meta-analysis (1977–2014), in: *Intelligence* 53, S. 145–513
- Rizzolatti, Giacomo/Sinigaglia, Corrado (2008): *Empathie und Spiegelneurone: Die biologische Basis des Mitgefühls*, Berlin
- Rost, Detlef H. (2013): *Handbuch Intelligenz*, Weinheim
- Salovey, Peter u. a., Hg. (2004): *Emotional intelligence, Key readings on the Mayer and Salovey model*, New York
- Spitzer, Manfred (2014): *Digitale Demenz: Wie wir uns und unsere Kinder um den Verstand bringen*, München
- Stern, William (1912): *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern* (Sonderabdruck aus: Bericht über den V. Kongress für experimentelle Psychologie/Berlin 1912), Leipzig
- (1918): *Person und Sache: System der philosophischen Weltanschauung. Zweiter Band: Die menschliche Persönlichkeit*, Leipzig
- (1938): *General psychology from a personalistic standpoint*, New York
- Thorndike, Edward L. (1920): Intelligence and its use, in: *Harper's Magazine* 140, S. 227–235
- Wirtz, Markus A., Hg. (1920): *Dorsch-Lexikon der Psychologie*, Bern
- Wittgenstein, Ludwig (2001): *Philosophische Untersuchungen*, Berlin
- Wongupparaj, Peera u. a. (2017): The Flynn effect for verbal and visuospatial short-term and working memory: A cross-temporal meta-analysis, in: *Intelligence* 64, S. 71–80
- Online-Quellen, Zugriff: 17.10.2020
- <https://www.vox.com/2015/7/2/11564066/elon-musk-backed-group-attempts-to-avert-judgment-day-with-ai-rules>
- <https://www.computerworld.com/article/2854997/stephen-hawking-says-ai-could-end-human-race.html>
- <https://www.duden.de/rechtschreibung/Intelligenz>
- <https://www.sciencedirect.com/science/article/pii/S0007681318301393?via%3Dihub>
- <https://www.zeit.de/2013/13/Intelligenzforschung-Zeitleiste>

Artificial Intelligence: Thoughts from a Psychologist

Michael J. Meitner

Zusammenfassung

In ihrem gegenwärtigen Entwicklungsstadium ist Künstliche Intelligenz (KI) noch weit davon entfernt, die Komplexität des menschlichen Gehirns zu erreichen. Der technische Fortschritt könnte jedoch zu einer KI führen, die als eine allgemeine Intelligenz die unsere in wichtigen Aspekten übertrifft. Ein entscheidender Schritt nach vorne könnte darin bestehen, dass KI ein Ausmaß an Diversität annimmt, wie wir es bereits aus unserer eigenen Neurologie her kennen. Für den weiteren Weg ist es wichtig, die Ziele im Blick zu behalten, die mit der Entwicklung von KI verbunden sind. Sich selbst überlassen, könnte KI zu einem verheerenden Wettbewerber und Kontrollfaktoren für Menschen werden. Dazu gestaltet, uns mit unseren Herausforderungen zu helfen, v. a., was die Umwelt betrifft, könnte KI aber auch einer menschlichen Entwicklung zuträglich sein.

Abstract

In its current state, Artificial Intelligence (AI) is still very far from reaching the complexity of the human brain. Technological progress, however, might bring about AI as a general intelligence surpassing our own in important aspects. One crucial step might be AI gaining an amount of diversity as we know it already from our neurology. On the way forward, it is important to consider the goals connected with the development of AI. Left to its own devices, AI might turn into a devastating competitor and controlling factor for humans. Designed to help us with our chal-

lenges, especially concerning the environment, AI might as well benefit humane development.

1. Introduction

Artificial intelligence, or AI as it is commonly referred to, is a suite of technologies that are poised to change the world as we know it. The concept of AI has been with us throughout antiquity in the mythologies of the Greeks and in early conceptions of automata. Early work in cybernetics and eventually neural networks brought this concept out of the realm of fantasy and into the modern world. In 1952, Marvin Minsky and Dean Edmonds succeeded in creating the world's first functional neural network machine, the Stochastic Neural Analog Reinforcement Calculator or SNARC¹. While this was no doubt an impressive achievement of the day, it did not really live up to the dreams of the ancients of a machine that would embody more human characteristics.

The holy grail of AI has always been to create a machine capable of generalized intelligence. In fact, the first known test of generalized intelligence in AI was posited by Alan Turing in his seminal paper "Computing Machinery and Intelligence"² and to this day is known as the Turing Test. The point of this test was to ascertain if a machine could fool a human into thinking it was conversing with an actual human being. When I was in graduate school, I had the pleasure to interact with ELIZA, a computer program written in 1966 by Joseph Weizenbaum to mimic the behavior of a Rogerian psychotherapist³ which some would argue was the first program to pass the Turing Test. While this contention remains controversial to some I can personally attest to the convincing nature of the program. However, I can also attest to the fact that it was quite easy to trip up this software and therefore destroy the illusion. It could never imagine new or novel situations and often answered any question that required creativity with a question of its own. Some historians of the internet age might say that ELIZA represented the first "bot", a software program that imitates the behavior of a human, as in participating in

¹ Minsky (1952).

² Turing (1950).

³ Weizenbaum (1966).

chatroom or IRC discussions. As most of us know today, bots have become far more sophisticated and for many of us they seem quite human when we interact with them.

However, there remains a massive disconnect between imitating a human and creating an artificial human brain. The human brain contains approximately 86 billion neurons and each neuron has on average 7,000 synaptic connections yielding nearly a quadrillion synapses⁴. In terms of simple computational power (measured in floating point operations per second or FLOPS) the human brain is estimated to be capable of approximately 1 exaFLOP (10^{18}).¹ Modern technology still falls short of this degree of raw computational power. The world's fastest supercomputer, Fujitsu for Japan's RIKEN Center for Computational Science supercomputer, has currently achieved 422 exaFLOPS⁵. However, it should be noted that important architectural aspects of the human brain are even further from the realm of possibility currently. Simply having the ability to do the same number of calculations over time does not mean that the arrangement of those neuronal units is in anyway similar to that of a human brain. Even in the case of AI modeling of *C. elegans*, a common worm that has only has 302 neurons, researchers are still refining the architecture of that model based on new electron microscopy data⁶. Therefore, the goal of a generalized intelligence instantiated in a computer is likely very far in the future. One possible technological development that may change this calculus is quantum computing but this still has significant challenges to overcome to become relevant to this discussion. Computational power (quantum or not) will certainly close the gap but this belies the fact that human brain is not simply the sum of its abilities to do raw computations.

In general, I would say that AI, in its current form, is in no way like the human brain even though AI researchers use architecture developed from observations of neuroanatomy. Modern AI is mostly focused on “narrow, shallow or weak AI” tasks such as finding patterns in our purchases and suggesting new ones based on these patterns. Even those AI's considered “broad, deep or strong AI” do not really approach the complexity of the human brain. Deep AI consists of numerous neural networks often hierarchically arranged that allow for deeper levels of ab-

⁴ Drachman (2005); Herculano-Houzel (2009).

⁵ Fulton (2020).

⁶ Cook et. al. (2019).

straction from the inputs in the model. In addition, deep AI techniques deal well with unstructured data and can analyze that data in an unsupervised fashion. These qualities have made deep learning techniques quite ubiquitous and they have been employed to tackle problems such as speech recognition and computer vision. Artificial general intelligence, on the other hand, will require substantial leaps in both hardware and software before this can be realized.

At this point I would like to compare and contrast the nature of artificial and human general intelligence as seen in table 1 below.

Artificial	Human
Infinite sensors	Limited sensors (can be augmented)
Infinite dimensions	Dimensionally challenged
Infinite data storage	Limited
Technology bound	Organism bound
Fairly stable goals (can be made to evolve)	Changing and evolving goals
Ever increasing processing speed	Speed mostly fixed
Replication generally yields copies (unless a genetic algorithm is used)	Replication yields neuro-diversity
Consciousness?	Multiple unconsciousness systems partially discovered by consciousness

Table 1: A comparison of artificial and human general intelligence

As is evident from table 1, artificial general intelligence holds much promise and will likely lead to the formation of an artificial superintelligence. Being able to surpass our limitations in data sensing, data storage (memory in humans) and in hyper-dimensional thinking at speed will allow AI's to make tractable those problems that have long eluded us. The goals of AI and the eventual architecture (and potential diversity of architecture) seem to be important turning points in our thinking about how we might make progress toward the creation of artificial general intelligence. Let us start by first turning our attention to diversity in AI.

2. Neuro-diversity and AI

The concept of neuro-diversity has been with us since 1998 and refers to the revelation that variation in the human brain is vast and while some variation may be detrimental other variations may represent significant strengths or improvements. In fact, I would go as far as saying that neurodiversity can in fact represent a competitive advantage. If true for humans, this surely would be true of diversity in AI as well. It has long been known that genetic algorithms (GAs) can be used to spawn novel architectures for neural networks that can be used to evaluate the degree of performance of its progeny on some fitness function or goal. This allows for competition between various forms of an AI algorithm and leads to better solutions to problems that the AI is tasked with. This represents some degree of neurodiversity in AI already, albeit a weak form of it, as unsuccessful progeny are “killed off” and therefore diversity is not maintained. Ideally, neuro-diverse AI systems would be persisted and alternate solutions could be investigated to allow for insights into divergent approaches that may help us to better define and build robust and resilient AIs in the future.

Ultimately the discussion of the concept of neurodiversity in the context of AI causes us to question our ideas about goals. Goals in AI must be made explicit in some way and often represent the most challenging aspect of creating a functional AI. For many AIs there are more than one goal that the algorithm is trying to maximize or balance amongst. However, all of these goals have a context and perspective. From a user’s perspective, a common goal might be increasing the relevance of information retrieved based on a query. From the companies’ perspective a similar goal might be user engagement. These differences in defining goals can have significant effects on the outputs of an AI. In fact, they define them. Variability in goal definition over time allows a model to adapt to changing system conditions.

As referenced in table 1, human goals seem to be ever changing and evolving as our understanding of the world progresses. This is especially true in the case of “wicked” problems. Wicked problems are those that defy simple solutions and are often comprised of multiple interacting systems. They are wicked because they are typically poorly understood, include contradictory information and are highly variable over time. Wicked problems do not have an optimal solution, rather they have temporary or partial solutions that are likely themselves to change over time. The changing nature of wicked problems and the large uncertainties in their

predictions mean we have to take an adaptive approach to the problem. Like wick-ed problems, adaptive problems are where the problem definition is mostly unknown. Adaptive problems often require the locus of control for solving the problems to be decentralized. Stakeholders become the focus rather than disciplinary experts and as we well know stakeholders often have a variety of perspectives on a problem. This is the type of diversity needed if we hope to be able to conceptualize the system properly. From that one might argue that this means that multiple AIs might be needed to focus on various specificities of a problem for a larger definition of the problem to occur.

3. AI in the environmental sciences

Many of the issues of our day are in fact adaptive problems, such as most of our problems. The World Economic Forum report titled “Harnessing Artificial Intelligence for the Earth”⁷ states that there are six priority action areas for addressing environmental issues: 1) climate change, 2) biodiversity and conservation, 3) healthy oceans, 4) water security, 5) clean air and 6) weather and disaster resilience. Each of these areas has a series of sub areas that AI could be applied to in order to create a more sustainable future. In the case of climate change, they refer to: clean power, smart transport options, sustainable production and consumption, sustainable land-use, smart cities and homes. AI can be applied to all these areas and in certain cases have the potential to transform these sectors. Consider a modern energy grid that can use AI to adapt to changing supply and demand, incorporate traditional power sources with clean energy and to make distributed energy possible at scale. This would seriously improve our ability to meet our climate change targets. As well, significant improvements in transportation, agriculture, and water management systems can also be realized by the application of AI technologies.

AI has already been applied to many environmental problems. Monitoring endangered species⁸, tracking diseases⁹, crop optimization¹⁰, smart buildings and

⁷ World Economic Forum (2018).

⁸ Dujon/Schofield (2019).

⁹ Wong et al. (2019).

¹⁰ Talaviya et al. (2020).

associated IoT to increase efficiency¹¹, predicting storms¹², and managing traffic¹³ are but a few of the many applications of AI in the environmental domain. In all these cases, AI offers us a method to deal with the massive degrees of complexity that represent these wicked environmental problems. This is made possible by the vast quantities of data that we are currently collecting to support decision making in these areas.

The world of “big data” has arrived and no technology is better poised to make use of this plethora of data than AI. In fact, without computer aided decision making, I would venture to guess that we would not be able to effectively navigate, understand or even utilize the amount of data that is currently available. AI, however, has a special relationship with big data and becomes better when provided with increasing data volumes. AI is especially good at detecting anomalies in massive data sets, determining the probabilities of future outcomes and it can recognize patterns that human cannot.

4. AI and Risk

The same report that was mentioned above also identifies six areas of risk for AI¹⁴. They are: performance, security, control, economic, social, and ethical. Performance risks refer to problems in deciphering the “black box” inner workings of an AI. Because we have little insight into what an AI is actually doing, we have difficulties in knowing if its performance is accurate or even desirable. Issues of model fit are also complicated by this. If an AI is inferring future trends based on historical records then we need to wonder if those records contain enough information to support such prediction. If we don’t know what an AI is doing internally then this problem is exacerbated.

Security risks, mentioned in this report, are also of concern. They reference “hackers” and the problems of bad actors manipulating algorithms to take control of them. This brings to light a more serious concern of who has control over these

¹¹ Panchalingam et al. (2019).

¹² McGovern et al. (2017).

¹³ Abduljabbar et al. (2019).

¹⁴ World Economic Forum (2018).

algorithms. Most AIs are in the hands of governments or large private sector companies. Neither of these has a great track record of acting for the social good. Private companies have a fiduciary duty to act in the best interests of and while they may make efforts to address social issues this will never be their primary concern. However, one could argue that a government's main interest is the public good but as we all know this can be perverted in service of other goals that do not in fact create nor maintain social good. Even if these actors had social good in mind, how is it defined? Would those actions taken by these actors result in increased social good? This is an open question and certainly needs more thought and discussion to determine how to fully define this risk.

Control risks are some of the most blown out of proportion but are also some of the most worrying. This is where common narratives of post apocalyptic worlds governed by intelligent machines that have decided that humans represent a threat come in. However, this does not really represent a credible threat because you would need an AI capable of general intelligence and we have already determined that the likelihood that this will materialize in my lifetime is remote at best. What is of more concern are AIs that have direct control of various systems that might make decisions that lead to unintended consequences. One example of this is the flash crash of the US stock market in 2010 which was likely caused by interaction of multiple AI bots all speed trading at the same time¹⁵.

Economic risks are also potentially significant for AI as it moves forward. Companies that do not have access to AI or the associated data to drive them run the risk of being out competed. This in turn creates the risk that the business landscape will continue to shrink, creating increased inequity of wealth distribution and consolidating power with a few multinational companies. This may lead to a circumstance where a few companies begin to exert more power over global progression.

Social risks of AI are often defined as adaptation to increased automation pressures created by increased use of AI. Job loss and increased unemployment are real possibilities in a world where AI takes over much of the work of running the systems that we rely on. Additionally, AI algorithms can potentially be biased against certain factions of society, underpinning historic social inequities. New inequities can also be created by AI as it fundamentally changes the sector with in which it is being applied. Take as an example the transportation sector where autonomous

¹⁵ Lauricella et al. (2010).

vehicles are poised to massively disrupt people's lives who rely on this sector for employment.

The last risk that this report discusses is ethical risks. What choices will an AI make? Will they be beneficial choices? What about fairness and human rights? Privacy concerns are also discussed here. While all of these risks are important and represent an excellent attempt to get us all thinking about how AI will shape our future, I believe that there is a significant omission in the risks associated with continued development and application of AI technology.

5. Additional risks

I would add two additional risks to this list; access to both data and the knowledge needed to make sense of it. Let us first tackle data access. AI does not represent a valuable technology without the data that drives it, and data is not generally freely available. Of course, there are open data sets but the vast majority of meaningful data being generated today is in the hands of private corporations or governments. In 2020, every minute of every day we collectively generate 500 hours YouTube video, WhatsApp users share 41,666,667 messages, Facebook users upload 147,000 photos, Instagram users post 34,722 stories, and TikTok is installed 2,704 times¹⁶. The amounts of data being generate currently is staggering and for the most part we create this data. It is estimated that in 2020 each person on earth generates 1.7 MB of data per second¹⁷. Because access to this proprietary data is in the hands of the few, and by all accounts, the powerful, we run the risk of increasing inequity in society. Not just in terms of wealth, which is certainly an issue worth discussing, but also in terms of access to the information being derived by various AIs. How are common people supposed to keep up when knowledge about our behaviour, actions, purchases, interests, beliefs, and values are being used to manipulate us? To control our purchases, our information feeds, our attention, our very lives. Something must be done to level the playing field. At a bare minimum we should have access to information about specifically how this data is being used to influence us.

¹⁶ Domo (2020).

¹⁷ Id. (2018).

Levelling the playing field however is not an easy task. There are many issues that need to be dealt with before we can hope to begin to bring us closer to balance. One of the first is the fact that this data is often privately owned. By agreeing to the licensing agreements (that honestly, we don't really have much of a choice about), we have given up our rights to this data (as per the individual agreements). Additionally, we might also be concerned about privacy. No one really wants their neighbor to have access to their search history. This later problem however, is a far more tractable problem. Data can always be anonymized and abstracted to hide individuals within the masses as is commonly done with census data. The real sticky wicket is the ownership issue. Companies will not give up this data without a fight. This data represents real value to these companies and access to these data sets is often sold to third party companies for a variety of reasons. If this data were freely accessible to all it would significantly alter the business model for many companies that specialize in this area. If this is not remedied however, we can expect the knowledge divide in society to grow and eventually this may in turn weaken the functioning of civil society in the future.

One possible solution to this is to consider making companies that supply services that are critical to civil discourse, public utilities and regulate them as such. This would ensure fair and equal access to these platforms that give citizens voice. No one can tell you that you can't have a phone and as well no one censors what you say when you are participating in a phone call. Why should digital communication services be any different? One argument would be that today's digital communication platforms are in the public sphere rather than a private communication between individuals but this simply changes the scope of the communication and who can see it. Currently our approach to this is one of censorship. We disallow those things that we find offensive and label it hate speech, striking any record of it from our collective discourse. To some this is seen as necessary to ensure a peaceful and equitable society, to others this is seen as top down control by those in power to limit personal freedoms. The real question about censorship is not whether we should to it but who is doing it? Who gets to decide what appropriate speech is? If you are in charge of this then I would imagine that you would be quite happy with the rules but others might think of you as intolerant. In my opinion speech should be protected unless it directly leads to action that is prohibited; violence, harassment, etc. Otherwise you have to decide what speech is acceptable and what is not and as history has taught us, this is a slippery slope. Once the precedent is set then

even if the previous government enacted censorship laws that we consider ethically correct, the next party in power could use this same power to rewrite the laws in their favor and impose restrictions on speech that may not be as ethically centered. Take for instance the case of the National Union of Students who in 1973 got racist speech banned at universities in England. This ban was supported at the time by an organization of Zionist students. For this seemed like a win but a few years later a different group of students was in power at the National Union of Students and they decided a Zionist speaker from speaking on campus because they now considered Zionism a form of racism. As you might have imagined, the group of Jewish students likely did not see how this might be turned against them as the leadership of that organization changed over time¹⁸.

If these companies are considered public utilities, we could also mandate that the proceeds of all analysis (knowledge) of our collective data should be freely available to everyone. This could come in the form of information dissemination and outreach on the part of the companies involved or it could mandate free and equal access to this data for the purposes of analysis. Both have their strengths and weakness but it seems to me that allowing companies to be the sole arbiter of what gets published is a bad idea. If we pursued the later idea then we would need to find ways to make these vast amounts of data available in real time. Additionally, there are numerous barriers to fair and equitable access to this data even if provided freely. Access to sufficient computing hardware and software is required for anyone to begin the process of data analysis of these massive data sets. This is certainly not equitably distributed either. As well the knowledge required to not only conduct such an analysis but to comprehend it as well.

This brings me to my second point regarding risks; access to education and the knowledge that it brings to the individual is critical for individuals to have sufficient skill and training to approach this analysis with rigor and accuracy. To some degree we are far closer to this goal than we are to the goal of equal access to the data itself. Online education has exploded over the years and many topics such as computer programming skills are currently freely available to those that have the inclination to pursue them. This does not mean that they have access to the best and brightest minds on the subject but they do at least have enough access to learn most of what would be required of an AI researcher today. This would allow many more minds to

¹⁸ Glasser (2020).

be focused on common problems that we face today but also to potentially uncover new and previously unknown ideas at a far greater rate. It is in this exploratory space that I see this type of citizen science as being most directly applicable. With more minds come more perspectives, potentially allowing us to see a greater degree of the underlying “Truth” of the world. This is certainly in line with the ideas presented earlier related to neuro-diversity.

A criticism of this approach might be that there is little control over the preparedness of individuals that seeks to undertake this type of work. However as one can plainly see this has always been the case. Even today not all researchers are considered equal. Some have tremendous knowledge and insights into the complexities of this undertaking and it is highly likely that contributions by these individuals would be of more import. The solution to this is as it always was. Peer reviewed publication practices can go a long way to maintaining a high standard when it comes to the quality of our collective scientific efforts. However, we also have to aware of the fact that the academic-industrial complex does not have exclusive license to seek the truth. Many minds of great importance do not get the chance in life to contribute to their full potential. Creating a strong program of citizen science, free and open data sources, and access to the knowledge required to pursue such endeavors is paramount for our society to move toward a collective vision of a future where discourse is alive and well, we share that which has the potential to collectively move us forward, and allow all voices to participate in the creation of said future. It is my great hope that we can find new ways to set our collective table in such a way that all leave nourished in mind, body and spirit.

6. Conclusions

AI is quickly redefining our world and if we continue along our current tack we will likely exacerbate social inequalities and eventually make a less stable world for our children. This is the challenge for the science of AI. Can it mature quickly enough to provide us with insights and abilities that may help us to create a more sustainable and equitable future for all? Climate change and associated global risks are the challenge of our time and human nature is likely the root cause of this dilemma. AI offers us the potential to turn the light of science on our interior nature and the ramifications that this has for our collective future and our future actions within it.

Make no mistake, trying to understand our collective behaviour, is the most wicked problem of all. Made even more so because we are the both the cause and solution to the problem. We are on the dance floor of our own perceptions, emotions and thoughts and we need to get on the balcony to be able to see the patterns that are emerging. AI offers us this vantage point. Granted we have a long way to go to improve the science of AI to allow us this potential but it exists none the less. I would hope that we could find a way to act in the collective good. To create a digital world where the rights to participate in society are inviolate, where access to data critical to said discourse is guaranteed, and diversity of perspective is the only requirement for entry.

Social networking companies must begin to think about the world they are allowing us to flow into. We are all on a river of time and the topography that underlies that river is the very nature of our digital (and physical, etc.) world that we have created to date. But just as topography yields to the bulldozer, our digital landscape is ours to remake. Let us start a discourse on this topic. Let all the world's peoples participate. We now have the tools to make this possible. It is world, but it is also one that must continue to improve if we are to hope engineer ourselves out of the current environmental trajectory.

Bibliography

- Abduljabbar, Rusul et al. (2019): Applications of artificial intelligence in transport: An overview, in: *Sustainability* 11/1, p. 189
- Brain performance in FLOPS (n.d.) [<https://aiimpacts.org/brain-performance-in-flops/>, retrieved 26.01.2021]
- Cook, Steven J. et al. (2019): Whole-animal connectomes of both *Caenorhabditis elegans* sexes, in: *Nature* 571/7763, pp. 63–71
- Crevier, Daniel (1993): *AI: the tumultuous history of the search for artificial intelligence*, New York
- Domo (25.01.2018): *Data never sleeps 6.0*. Domo.com [https://www.domo.com/assets/downloads/18_domo_data-never-sleeps-6+verticals.pdf, retrieved 02.02.2021]
- Domo (25.01.2020): *Data never sleeps 8.0*. Domo.com [<https://www.domo.com/learn/data-never-sleeps-8>, retrieved 02.02.2021]
- Drachman, David A. (2005): Do we have brain to spare?, in: *Neurology* 64/12, [<https://doi.org/10.1212/01.WNL.0000166914.38327.BB>]

- Dujon, Antoine M./Schofield, Gail (2019): Importance of machine learning for enhancing ecological studies using information-rich imagery, in: *Endangered Species Research* 39, pp. 91–104
- Fulton, Scott (18.11.2020) *Top500: Japan's Fugaku Still the World's Fastest Supercomputer* [<https://www.datacenterknowledge.com/supercomputers/top500-japan-s-fugaku-still-world-s-fastest-supercomputer>, retrieved 26.01.2021]
- Glasser, Ira (2020, October 4): *How Freedom of Speech Protects You from Rulers like Trump* [<https://www.msn.com/en-us/news/opinion/aclu-hero-ira-glasser-how-freedom-of-speech-protects-you-from-rulers-like-trump/ar-BB19GBC5>, retrieved 26.01.2021]
- Herculano-Houzel, Suzana (2009): The human brain in numbers: a linearly scaled-up primate brain, in: *Frontiers in human neuroscience* 3, p. 31
- Lauricella, Tom et al. (October 2, 2010): How a Trading Algorithm Went Awry, in: *The Wall Street Journal* [<https://www.wsj.com/articles/SB10001424052748704029304575526390131916792>, retrieved 26.01.2021]
- McGovern, Amy et al. (2017): Using artificial intelligence to improve real-time decision-making for high-impact weather, in: *Bulletin of the American Meteorological Society* 98/10, pp. 2073–2090
- Minsky, Marvin (1952): A neural-analogue calculator based upon a probability model of reinforcement, in: *Harvard University Psychological Laboratories*, Cambridge, Massachusetts
- Panchalingam, Rav et al. (2019): A state-of-the-art review on artificial intelligence for Smart Buildings, in: *Intelligent Buildings International*, pp. 1–24
- Talaviya, Tanha et al. (2020): Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides, in: *Artificial Intelligence in Agriculture* 4 (2020), pp. 58–73
- Turing, Alan (2009): Computing machinery and intelligence, in: Robert Epstein et al., eds., *Parsing the turing test. Philosophical and Methodological Issues in the Quest for the Thinking Computer*, Dordrecht, pp. 23–65
- Weizenbaum, Joseph (1966): ELIZA—a computer program for the study of natural language communication between man and machine, in: *Communications of the ACM* 9/1, pp. 36–45
- Wong, Zoie S. Y. et al. (2019): Artificial intelligence for infectious disease big data analytics, in: *Infection, disease & health* 24/1, pp. 44–48
- World Economic Forum (2018): *Harnessing Artificial Intelligence for the Earth*, Geneva

Nature as a Work of Art?

translated into English by Kevin Fencil

Marion Friedrich

Zusammenfassung

Unser Verständnis (unser „Begriff“) der Natur ändert sich mit zunehmenden technischen Möglichkeiten: Schon lange betrachten wir sie nicht mehr nur und bilden sie ab – wir zerlegen sie in Einzelteile und bemühen uns, sie uns zugänglich zu machen – das, was unsere Sinnesorgane nicht funktionsintrinsisch wahrnehmen können, amplifizieren und verändern wir. Auf diese Art schaffen wir u. a. Synästhesien: „Sounds of decay“, das Forschungsprojekt mit künstlerischen Ambitionen von Cat Hope¹ ist ein Beispiel. Tod, Zerfall von lebender Materie, sichtbar mit und ohne Hilfsmittel, wird hörbar gemacht.

Auf der anderen Seite verschieben wir unsere „natürlichen“ Grenzen als Menschen durch die Implementierung künstlicher (technischer) Komponenten in den lebenden Organismus mehr und mehr. Jeder Versuch, Natur (über die Erkenntnisse der Einzelwissenschaften hinaus) begrifflich zu fassen, führt in letzter Konsequenz zur Frage nach der Natur des Menschen zurück: Diese Hürde zeigt sich zunächst unüberwindbar. In jedem Bemühen, Natur zu objektivieren, setzen wir sie in Relation zu uns als Subjekte. Insbesondere die Möglichkeiten der modernen Technik stellen uns erneut vor die Herausforderung, Stellung zu uns als natürliche, in der Natur befindliche (und von ihr abhängige) Lebewesen zu beziehen – und dies unabhängig davon, welches Gegensatzpaar wir in dem Bestreben, uns einen Naturbegriff zu bilden, auch konstruieren (Natürlichkeit – Nicht-Natürlichkeit; das Ent-

¹ Sounds of Decay (2013). Sound Installation. Premiered at SemiPermeable at the Powerhouse Museum, Sydney, as part of the International Symposium of Electronic Art. Made in collaboration with Rob Muir, as part of an artist residency at SymbioticA lab.

standene – das Gemachte; das Normale – das Abnormale; das Authentische – das Künstliche; Gesetzmäßigkeit – Willkürlichkeit).

Je nachdem, wie wir Natur definieren, gelangen wir zu gänzlich unterschiedlichen Schlüssen in Bezug auf die Legitimität des (naturgemäßen?) Bestrebens, eben jene zu überwinden bzw. die uns durch unser Menschsein natürlich gegebenen Grenzen zu erweitern. „The primary political and philosophical issue of the next century will be the definition of who we are“, schrieb Ray Kurzweil². Noch nie schien die Kluft zwischen dem, was wir tun können (was technisch realisierbar ist), und dem, was wir tun möchten (was wir für moralisch richtig, für unanfechtbar halten), größer: Doch wie können wir entscheiden, ob beispielsweise Moral-Enhancement „richtig“ im Sinne von „natürlich“ genug (= moralisch korrekt) ist, wenn wir uns noch nicht einmal darüber klar sind, was überhaupt die „Natur des Menschen“ ausmacht?

Dass wir in die Natur eingreifen, sie zu verstehen und verändern suchen, ist allgemein eine (moralisch) akzeptierte Tatsache. Seit jeher haben wir unsere Umwelt (die „Natur“) durch unser Eingreifen zu Gunsten der menschlichen Spezies manipuliert. Heute verfügen wir über die technischen Möglichkeiten, in die Natur des Menschen selbst einzugreifen – ohne sie bislang ausreichend definiert zu haben: Mittels Enhancement-Methodik nehmen wir Einfluss auf Kognitionen, Emotionen und sogar Wahrnehmungsmodalitäten.

Es ist dringend erforderlich, dass wir uns darüber verständigen, „welche Art und welches Ausmaß des Eingreifens gut ist“³, auch, wenn der Versuch, eine einheitliche „Naturphilosophie“ und sei es nur durch das Erreichen eines Konsens in Hinblick auf einen Naturbegriff, zu betreiben, utopisch wirkt.

Es war kein allzu langer Weg von der Feststellung, dass Kunst konzentrierte Natur sei⁴ über hybride Kunstskulpturen bis hin zu „Sounds of decay“, dem Projekt von Cat Hope, die als Mitglied der australischen interdisziplinären Forscher-/Künstlergruppe SymbioticA, den Zerfall lebender Materie vertonte: Das Zellsterben einer in einem Luftentfeuchter gehaltenen, austrocknenden Riesenkröte wird in Musik umgewandelt. Auf diese Art machen wir Natur zum Kunstprodukt, machen sie uns für alle unsere Sinnesmodalitäten zugänglich, erweitern zugleich mit Neuroprothesen unsere Sinneskapazitäten.

² Kurzweil (2000).

³ Siep (1999).

⁴ Balzac (2002).

In dieser paradoxen Situation, in der die einen daran arbeiten, moralisches Urteilsvermögen (mit konsequentem Handeln) durch Eingriffe in die „menschliche Natur“ zu stärken, und die anderen „natürliche“ Prozesse pervertieren, um sie, auch unter Einbezug quasi synästhetischer Empfindungen, zu ästhetischen Objekten zu stilisieren, sich ihrer eigenen Natur damit entfremdend, wird Natur im doppelten Sinne zu einem Kunstprodukt: Natur wird zu konzentrierter Kunst.

Abstract

Our understanding of the term “nature” has changed as our technological prowess has grown. Whereas nature used to be something we observed and created representations of, it is now something we attempt to make accessible by disassembling into tiny pieces, amplifying and modifying that which our senses cannot otherwise perceive. In the process, we create experiences that are in part synesthetic. “Sounds of Decay,” in which death and the decay of living matter are made audible, is an example of this. Parallel to this, we’ve embarked upon an attempt to enhance our own “natural” abilities to perceive and perform by implanting artificial – technological – components in our bodies.

Findings within the individual branches of science notwithstanding, any attempt to make nature definable ultimately and inevitably leads to questions about the nature of man himself. Whenever we try to objectify nature, we do so relative to ourselves, the subjects. Regardless of which pair of opposites we choose (natural/not natural; that which originated/that which was made; the normal/the abnormal; the authentic/the synthetic; that which conforms to rules/that which is random), the possibilities afforded by new technologies present us with the challenge of constructing – or construing – a concept of nature and of taking a stance on what it means to be a human being, both of and dependent on nature. Depending on how we define nature, we arrive at widely varying conclusions about the legitimacy of our (natural?) efforts to overcome it, to go beyond the boundaries it has set out for us as humans.

“The primary political and philosophical issue of the next century will be the definition of who we are,” wrote Ray Kurzweil in the year 2000. The gap between what we are able to do (i.e. what was technologically possible) and what we want to do (i.e. what we considered morally defensible) has never seemed larger. How are

we to decide whether moral enhancement, for example, is “morally correct” in the sense of being “natural” enough when it isn’t even clear to us what the essence of the “nature of man” is?

That we intervene in nature and seek to understand and change it is a (morally) accepted fact. Human beings have been modifying their environment (nature) to suit their purposes since time immemorial. Today, we possess the technological savvy to manipulate the nature of man itself without having sufficiently defined the “nature of man.” Various methods of enhancement allows us to influence cognition, emotion and even modes of perception.

It is absolutely crucial that we reach a consensus about “what type and what degree of intervention is good” (Siep 1999), no matter how utopic the attempt to come up with a philosophy of nature we all can agree upon, or at the very least to reach a consensus regarding the definition of nature, may seem.

The path leading from the observation that art is concentrated nature (Honoré de Balzac, 1799–1850) via hybrid art sculpture to “Sounds of Decay” (a project by Cat Hope, who is a member of SymbioticA, an interdisciplinary Australian group of researchers and artists) is a short one. The dying cells of a cane toad dehydrating in a desiccator are transformed into music. Nature becomes a synthetic product, something we can experience with our senses, while our senses are enhanced with neural prostheses.

In this paradoxical state of affairs, in which one group strives to improve our ability to pass moral judgment (and act upon this judgment) by intervening in “human nature” while another group perverts “natural” processes and utilizes quasi synesthetic perception in order to create aesthetic objects that bear little relation to their natural selves, nature becomes not only a synthetic product but concentrated art.

1. Nature as a Work of Art

Suppose I arrived on Earth as a visitor from another planet and was taken on a sightseeing tour. If we first visited the river Lech as it runs its “natural” course⁵ and afterwards an artfully placed pond in somebody’s garden, I’d hardly be able to tell

⁵ Cf. Soentgen (2014).

which of the two is “natural” and which has been “made” – at least not without some sort of prior knowledge. I would see water, in varying degrees of motion, surrounded by a landscape – even the concrete bed of the pond is ultimately made out of “natural” materials⁶.

In order to differentiate nature from art, I have to apply the knowledge I have at my disposal about how the water came to be and about the rules and laws that govern its flow. In doing so, I would be introducing myself as an observer into the equation. What I could say is: If, by applying the methods available to me as a scientist, I can observe or at least assume the existence of some sort of adherence to a natural law, then the thing I am observing is nature. Water, for example, wants to flow downhill. Even a hill of sand piled up by a child can cause water to flow downhill, assuming there is enough mass to propel it forward.

Art seems to adhere to laws that can hardly be comprehended “objectively.” In everyday speech, art is thought of as something that has been made. Yet, is a bird’s nest *nature* or *art*? Whether we think the landscaped waterfall in our neighbor’s garden is pretty or aesthetically pleasing seems at first to depend on purely subjective preferences. Nonetheless, even our aesthetic sensibilities adhere to certain rules – both in art and in nature. Beauty, it turns out, is measurable. We consider a face to be beautiful, for example, when its features stand in a certain proportion to one another⁷. Both symmetry and “proximity to the average” (or “familiarity”) determine whether or not we find a face attractive in our culture⁸.

As a human being subject to the laws of nature, it is impossible for me to step outside myself and exempt myself from all the implicit rules that influence how I perceive things. Because of this, I can only perceive everything the world has to offer as a work of art. What qualifies as *nature* outside of my field of perception is impossible to say, given that man himself is a work of art.

⁶ Clearly, we can’t create “new” materials. All we can do is use existing “natural” materials and combine them, transform them or change their ratio to one another. At the microscopic level, their components remain the same.

⁷ When we emphasize certain features (when we elevate the familiar to the extraordinary, for example, by giving a face larger than average eyes), we create beauty, consciously turning “nature” into art.

⁸ Cf. <http://www.zib.de/deuffhard/pub/Schoenheit.pdf>, read, like all other internet resources referred to in this paper, retrieved 10.02.2021.

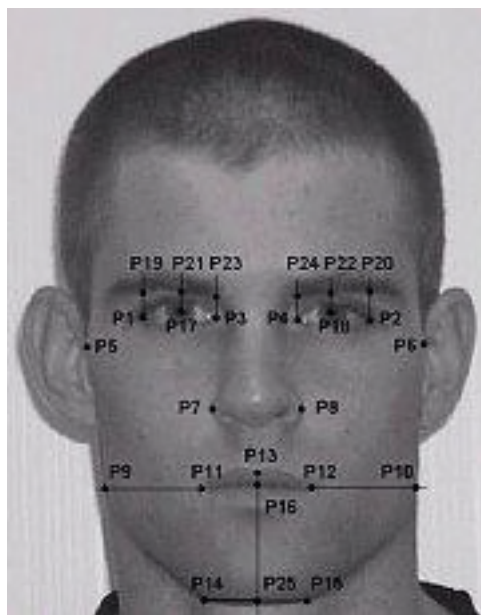


Fig. 1: Measuring beauty⁹.

Does this mean the terms “nature” and “art” are freely interchangeable – and thus meaningless? Does it make sense to take a philosophical view of nature at all when nature *as such* does not and cannot exist for us because it is, just like art, purely a contrivance of the human spirit?

I believe that today, more than ever, we must set ourselves to the task of uniting epistemology, hermeneutics and other philosophic semi-disciplines so that guidelines in terms of practical ethics can be established.

This paper will attempt to show that both nature and art are abstract terms that are, at first glance, indistinguishable from one another. After all, each time we observe nature, our observations are colored by our own self-image. Only when we begin to see ourselves as works of art does the opportunity to con-

⁹ Picture source: © Nicole Koehler; Koehler, N. et al.: The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proceedings of the Royal Society B* 271, Suppl 4, 2004, Fig. 1 (Section).

sciously plan our further evolution arise. There is no way of “returning to nature” as far as I can see.

Our understanding of the term “nature” has changed as our technological prowess has grown. Whereas nature used to be something we observed and created representations of, it is now something we attempt to make accessible by disassembling into tiny pieces, amplifying and modifying that which our senses cannot otherwise perceive. In the process, we create experiences that are in part synesthetic. One example of this is “Sounds of Decay,” a project by Cat Hope, a member of an interdisciplinary Australian group of researchers and artists called SymbioticA¹⁰. In it, death and the decay of living matter are made audible.

Parallel to this, we’ve embarked upon an attempt to enhance our own “natural” abilities to perceive and perform by implanting artificial – technological – components in our bodies.

Findings within the individual branches of science notwithstanding, any attempt to make nature definable ultimately and inevitably leads to questions about the nature of man himself. Whenever we try to objectify nature, we do so relative to ourselves, the subjects. Regardless of which pair of opposites we choose (natural/not natural; that which has originated/that which was made; the normal/the abnormal; the authentic/the synthetic; that which conforms to rules/that which is random), the possibilities afforded by new technologies present us with the challenge of constructing – or construing – a concept of nature and of taking a stance on what it means to be a human being, both of and dependent on nature. Depending on how we define nature, we arrive at widely varying conclusions about the legitimacy of our (natural?) efforts to overcome it, to go beyond the boundaries it has set out for us as humans.

“The primary political and philosophical issue of the next century will be the definition of who we are,” wrote Ray Kurzweil in the year 2000¹¹. The gap between what we are able to do (i.e. what is technologically possible) and what we want to do (i.e. what is considered morally defensible) has never seemed larger. How are we to decide whether moral enhancement, for example, is “morally correct” in the sense of being “natural” enough when it isn’t even clear to us what the essence of the “nature of man” is?

¹⁰ SymbioticA (2013).

¹¹ Kurzweil (2000).

That we intervene in nature and seek to understand and change it is a (morally) accepted fact. Human beings have been modifying their environment (nature) to suit their purposes since time immemorial. Today, we possess the technological savvy to manipulate the nature of man itself without having sufficiently defined the “nature of man.” Various methods of enhancement allow us to influence cognition, emotion and even modes of perception.

It is absolutely crucial that we reach a consensus about “what type and what degree of intervention is good”¹² no matter how utopic the attempt to come up with a philosophy of nature we all can agree upon, or at the very least to reach a consensus regarding the definition of nature, may seem.

The path leading from the observation that art is concentrated nature¹³ via hybrid art sculpture to “Sounds of Decay” is a short one. The dying cells of a cane toad dehydrating in a desiccator are transformed into music. Nature becomes a synthetic product, something we can experience with our senses, while our senses are enhanced with neural prostheses.

What motivates a person to end another living being’s life with the goal of creating a work of art, even if it is only the life of a toad? Where does the idea arise, not just of observing the process of decay but of creating out of it a “melody of death?” Is it (scientific) curiosity, the artistic muse or anger¹⁴ at the toads that seem to have descended on Australia like a plague? Whatever the motive may be, I would like to state for the record that I consider killing a toad and audibilizing its decay an act of aggression¹⁵.

In this paradoxical state of affairs, in which one group strives to improve our ability to pass moral judgment (and act upon this judgment) by intervening in “human nature” while another group perverts “natural” processes and utilizes *quasi*-synesthetic perception in order to create aesthetic objects that bear little relation to their natural selves, nature becomes not only a synthetic product, but concentrated art.

¹² Siep (1999).

¹³ Honoré de Balzac (1799–1850), Source: Balzac: *Illusions perdues* (1837–1843).

¹⁴ In general, artists are more likely to occupy themselves with emotions than scientists.

¹⁵ For the purposes of this paper, I would define aggressive behavior as an act aimed at adversely affecting the well-being of another living creature or, at the least, knowingly accepting that this will happen, independent of motive or intent.

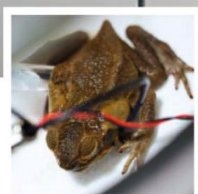
I'd like to explain, from a humanistic perspective, why finding a "way back" (to nature) seems impossible to me. I assume that in formulating a contemporary philosophy of nature, we are allowed to define ourselves as beings that occur within nature and to declare our (environmental and social) natural habitat a work of art, which we may sculpt according to ethical guidelines to be defined. A nature to which we might wish to return, assuming we found it to be a pleasant place, is therefore inaccessible to us. We have various strategies at our disposal to help us evolve as humans. We could decide, for example, to exert direct influence over our emotions, using moral enhancement or transindividual consciousness to weaken them or even negate their influence in favor of rational decision-making processes. Or we could concentrate on sculpting our natural and social habitat to optimally suit the needs of our species. Taking the humanistic view of man as a social being, this would entail reaching fair compromises between the needs of each individual. In such a scenario, primal emotions such as fear, anger, sorrow and disgust, which forcefully compel us to act and which often lead to destructive behavioral strategies, could be interpreted as "status indicators." This would allow us to concentrate on the optimal use and further enhancement of our higher cortical functions – our reason.

2. Sounds of Decay

In the booklet accompanying her project, "Semipermeable +" from 2013, Symbiotica describes the musician Cat Hope's installation as a concept rather than a finished work:

From life, to un-life: a dead cane toad makes music/albeit very, very quietly. It's a sort of fuzzy, low, soft wound, usually inaudible, but processed for the human ear in Cat Hope's Sound of Decay (2013). The corpse itself is a vaguely disturbing form almost obscured by the condensation that lines the walls of its sealed glass container, and as such evokes death and kind of elongated timescale more sure, perhaps, than a clear view.

RealTime Arts on Cat Hope's project "Sounds of decay"



Cat Hope is a composer, sound artist, performer and academic. She is a classically trained (baroque) self-taught vocalist and experimental noise bass player with an active performance profile as a soloist and in music groups. She is the director of the internationally recognized music group, *Recibel* and has toured internationally as a noise artist. Her installations have been shown at ECA (Japan, Tallin), Liquid Architecture, the *Totally Huge* (New Music Festival), The Perth International Arts Festival and in Singapore, USA, Japan and Ireland. Cat is currently a researcher at the Western Australian Academy of Performing Arts at Edith Cowan University. This work is conceived by Cat Hope and Rob Huie, as the *High-Voice Sound Collective*, with funding from an Edith Cowan University Faculty Research Grant. The computer programming is by Stuart James, based on a concept by Bruce Murphy. www.sounddecay.com

Assisted by the State of Western Australia through the Department of Culture and the Arts.

Fig. 2: *Semipermeable+*, *Symbiotica*, 2013, p. 26, p. 27¹⁶.

¹⁶ Credit to photographer Ian Hobbs, 2013, artwork in the photograph Cat Hope.

3. Back to (human) nature?

From a humanistic perspective, man is good by his very nature. We are evolving individuals who strive to improve, wrote Andrés Sánchez Bodas in his *Manifiesto humanístico*¹⁷, willing to do another a good turn once we have realized that we, too, will be helped when we need it¹⁸. Attempting to understand man as an individual outside the context of social relationships would appear to be as fruitless an endeavor as trying to approach *nature* as something that we ourselves are not part of and have no relation to.

Martin Buber's description of the direct (and natural) relationship between the "I" and the "thou" in his work *Das dialogische Prinzip* sounds modern because, among other reasons, it appears to account for the findings of neuropsychological research into empathy and attachment behavior.

In this context, I understand aggression to be an act that harms not only another person, but me as well (because I am in a constant state of relationship to others). Aggression is a form of self-harm.

This is very much the opposite of the fundamental assumption of humanism I wish to promote, in which life strives *by its very nature* to remain intact, to maintain its physical and mental integrity. This leads me to conclude that aggressive behavior is a result of a faulty learning process and does not necessarily reflect human nature; rather, it countermands it. Every creature strives to remain intact and to preserve life, choosing to extinguish it only when placed in a life-threatening situation itself or in order to optimize the skills that help it remain alive¹⁹.

We humans would seem to be an exception. No longer caught up in the battle for survival, we do not kill out of fear, but for scientific purposes, curiosity, revenge or pleasure. In this sense, death can be seen as the final consequence of a behavior aimed at hindering another living being in its pursuit of well-being.

¹⁷ Bodas (2013).

¹⁸ Carl Roger's protégé Marshall Rosenberg put Roger's theoretical postulate into practice with his process of Nonviolent Communication.

¹⁹ Such as with housecats and other predatory cats that practice chasing mice or dismembering their prey.

As such, aggression isn't a feeling, but a behavioral strategy applied to satisfy unfulfilled needs which are the root cause of noticeable emotions. We human beings are in many cases no longer in touch with/no longer understand our own "nature."

If nature, and particularly the nature of man, is not objectively accessible to us – if we cannot comprehend the nature of nature – then this positive view of the nature of man seems to me to be no less plausible than any other²⁰. I therefore decide to believe in Carl Rogers' actualizing tendency²¹, which states that each individual strives to realize his or her own potential, in part by forming a congruous self-concept based on everything the organism as a whole has experienced both unconsciously and consciously²². In this sense, aggressive behavior isn't part of human nature, but rather an indication of our increasing *estrangement* from our nature.

Yet, how should we propose to return to our natural state when we cannot objectively determine what it is? When, whatever it may be, it is incapable of standing solitarily and discretely alongside that which we have learned and which has shaped us? When we can't even assume agreement upon one of the various "concepts of man" available to us?

In asking these questions, we eventually come to realize that there can be no direct way "back to nature." On the contrary, the challenge we face is of setting a course for the further evolution – the *progress* – of mankind and figuring out how we can influence that.

The discussion I'd like to spark is whether we can control our aggression²³ by accessing our innate ability to empathize. If strategies of ethical behavior are to be derived from what Theodor Lipps referred to as "*Einfühlung*"²⁴, certain environmental factors must be in place. When I say "ethical behavior," I mean behavior that does not pur-

²⁰ No theory of the nature of nature is verifiable because every attempt at empirical examination constitutes an intervention in or manipulation of that which is being examined. Obtaining objective findings would appear to be impossible; one cannot simply be aware, one must be aware "of something."

²¹ As opposed to the death drive (Todestrieb) assumed in classical psychoanalytic studies.

²² Rogers (1994).

²³ I would submit that every act of aggression is a superfluous act.

²⁴ *Einfühlung* was Theodor Lipps' term for empathy, i.e. the ability to transcend the "I" in order to participate subjectively in that which another person experiences. That this isn't only a conscious process, but primarily a pre-conscious and thus involuntary process has been confirmed by countless studies, some of which are referred to in this paper, cf. Lipps (2009).

posely harm another living being, assuming all living beings are of equal value. With this “negative” definition as a starting point (ethical behavior = removal of aggression), we can set up a “positive” guideline that would go something like this: Ethical behavior is behavior (usually a compromise) that contributes to the well-being of all involved.

If we do not succeed in ensuring that living conditions are such that all living creatures can satisfy their needs (and I assume there is no way “back to nature,” because we’ve gotten so accustomed to reducing abstract terms like “nature” to the superficial that it hardly occurs to us how absurd any attempt at “re-naturalization” would be), then the only (?) alternative left to us is to influence evolution in such a way that its result is ethically acting beings adhering to the principles of reason. This in turn would assume an ability to switch off emotions that lead to aggressive behavior. Which emotions might those be?

In an ideal world – so says humanistic thought – emotions are the result of fulfilled or unfulfilled needs, the purpose of which is to indicate those needs to us. We have a wide variety of strategies at our disposal for satisfying our needs. In the real world, we have forgotten how to differentiate between needs, emotions and courses of action.

In everyday speech, aggression is often conflated with emotion. The Oxford Dictionary defines aggression as “feelings of anger or antipathy resulting in hostile or violent behavior; readiness to attack or confront.”²⁵ The word has its origins in the Latin *ad* (towards) and *gradi* (proceed, walk). The connotation of an impending attack has been present in the word since the early 17th century.

I would like to trace aggression (aggressive behavior) back to emotions or, better yet, back to the unfulfilled needs that lie at the root of those emotions in order to better get at the question of how aggressive behavior can be modified and/or prevented. I would prefer, in doing so, not to get into a discussion of whether this is a desirable goal, and I don’t wish to delve into methods of operant conditioning such as punishing aggressive behavior. Psychopaths (violent criminals with dissociative personality traits who harm other living creatures seemingly only for their own pleasure²⁶), for example, are less responsive to pain inflicted on them as a con-

²⁵ <http://www.oxforddictionaries.com/definition/english/aggression>, retrieved 19.07.2021.

²⁶ The ICD-10 defines a dissocial personality disorder as follows:

- A disorder characterized by a pervasive pattern of disregard for and violation of the rights of others that is manifested in childhood or early adolescence (adapted from DSM-IV: Diagnostic and Statistical Manual of Mental Disorders).

sequence of their behavior than comparable groups²⁷. I would like to focus on prevention and, as such, on the design of human nature.

From a humanistic perspective, it is crucial to differentiate between behavioral strategies, feelings and needs in order to be able recognize consciously and with self-empathy one's own (un)met needs and the emotions (symbolized affects, cognitively labeled sensations) that result from them as they arise later on. Marshall Rosenberg's life work consists, among other things, of using non-violent communication to anchor in the consciousness of all humans the idea that every form of violence is a tragic expression of an unmet need.

Being conscious of this allows us to deviate from our automated behavioral strategies and to search for alternatives. In his master's thesis from 2005, Simon Beck²⁸ provides us with a simple set of directions for transforming the violence that results from needs calling out to be met, which often manifests itself as aggression:

*Transforming the Pain of Unmet Needs to the Beauty of Needs*²⁹

1. Acknowledge the stimulus, the neutral observation that you are responding to. Be specific and concrete, describing the precise stimulus for your feelings.
2. Acknowledge your reaction. There are 3 steps to transforming the "jackal":
 - a. recognition/naming the thought or message; b. embracing or "enjoying the jackal show". Allow any reactions, judgments, anger, etc. to come into

-
- A personality disorder whose essential feature is a pervasive pattern of disregard for, and violation of, the rights of others that begins in childhood or early adolescence and continues into adulthood. The individual must be at least age 18 and must have a history of some symptoms of conduct disorder before age 15 (from DSM-IV, 1994).
 - Personality disorder characterized by conflict with others, low frustration tolerance, inadequate conscience development, and rejection of authority and discipline.
 - Personality disorder whose essential feature is a pervasive pattern of disregard for, and violation of, the rights of others through aggressive, antisocial behavior, without remorse or loyalty to anyone.

²⁷ Scheider (1923).

²⁸ http://www.cnvc.org/sites/cnvc.org/files/NVC_Research_Files/NVC%20Research/Beck-Developing_NVC_Integral_Approach.pdf, retrieved 14.02.2021.

²⁹ Gonzales (2004).

your awareness, and express it to yourself silently, out loud, or *best written* for clarity. And c. differentiation from the jackal thinking. In this part you can use a phrase like, “I am telling myself ([...] the jackal message.)” Go through this process with each message.

3. What feelings arise in you? Notice/feel authentic feelings.
4. What are the unmet needs that give rise to the feelings? Stay with feelings and needs.
5. Stay with steps 3 and 4 until you have given yourself sufficient empathy. Allow yourself to “be with” the feelings of pain of your unmet needs. This is the mourning/grieving stage.
6. Stay in the mourning stage until you have felt a shift in your feelings. You will usually feel some relief and/or relaxing of feeling.
7. Very often what can occur at this stage is a re-stimulation of jackal thinking. If this happens, empathize with the feelings and needs behind this jackal and go back to mourning, or “being with” feelings and needs.
8. Now focus on the need itself. Not the unmet need, but the “beauty of the need.” Sense/feel the positive value, the inherent vision of why this need is important to you. Allow yourself to immerse your awareness on this aspect.
9. Notice any request you may have of yourself. What action(s), internal or external do you want to take to meet any need(s) that is/are present.
10. This entire process is not a linear, but rather a dynamic, organic process. You will probably move from one dimension to another, staying focused on what is alive.
11. The three qualities that come from this process are: *clarity*, *compassion* for self, and *empowerment* to move forward in deep self-connection and meeting needs.

4. Emotional Control?

B.F. Skinner has pointed out that suggestions that human behavior be controlled tend to elicit negative emotional reactions³⁰. Even in cases where behavioral control

³⁰ Cf. Skinner (2011), p. 3.

might contribute to enabling humanity to live together in mutual respect and recognition of one another's needs, a majority would probably refuse such measures. They might fear being manipulated or may interpret such measures as "curtailing their freedom." They could see them as a breach of human rights (do we have a natural right to behave aggressively?) or simply of their right to choose to behave wrongly if they wish. I define the choice of behaving wrongly as the freedom to behave in a way that is knowingly harmful to other living creatures (that violates their physical or psychological integrity), despite the fact that alternative ways of behaving which do not lead to harm exist. This is not to say that it is always intuitively clear which non-harmful alternatives are available.

I'd like to suggest that a guiding tenet of humanistic ethics be the idea that *in an ideal world*, a fair compromise between the needs of all participants can always be reached.

This approach to ethics is a utopian one, because it assumes that we no longer concentrate on the questions of how we will *reactively* behave, but how we can *actively* foster the conditions that allow ethical quandaries (or pseudo-quandaries) to be prevented³¹.

How can we create a world in which it is possible to distance ourselves from our perceived emotions, both noticed and unnoticed, in order to reach compromises driven by reason? How can we control our emotions (and the impulses they give rise to)?

Can we "turn off" needs so as to prevent emotions which are considered negative and which lead to aggressive behavior? I don't think we can. Needs, as I see them,

³¹ I am referring here to thought experiments along the lines of the so-called Kohlberg Dilemmas. For example: In Europe, a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to make. He paid \$400 for the radium and charged \$4,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money and tried every legal means, but he could only get together about \$2,000, which is half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So, having tried every legal means, Heinz gets desperate and considers breaking into the man's store to steal the drug for his wife. Other examples can be found at: <http://examples.yourdictionary.com/ethical-dilemma-examples.html>, retrieved 14.02.2021.

are a natural state of a living organism. What we can do is to recognize them for the life-sustaining functions that they are³² and, by *equitably* distributing the resources we have at our disposal, satisfy them before they lead to violence.

Can we manipulate the character of the perceived affects³³ that are emotions so that their influence on actions is weakened or even eliminated? Here, several options are open to us:

1. We could proceed indirectly, by satisfying our needs, so that negative emotions are counteracted while positive, reactive emotions (pleasure) are fostered. Here is Marshall Rosenberg's exemplary list of needs and the emotions that result from them:

Eight Primary Emotions:

- Sadness: grief, sorrow, gloom, melancholy, despair, loneliness, and depression.
- Anger: fury, outrage, wrath, irritability, hostility, resentment and violence.
- Fear: anxiety, apprehension, nervousness, dread, fright, and panic.
- Joy: enjoyment, happiness, relief, bliss, delight, pride, thrill, and ecstasy.
- Interest: acceptance, friendliness, trust, kindness, affection, love, and devotion.
- Surprise: shock, astonishment, amazement, astound, and wonder.
- Disgust: contempt, disdain, scorn, aversion, distaste, and revulsion.
- Shame: guilt, embarrassment, chagrin, remorse, regret, and contrition.

All other emotions are made up by combining these basic eight emotions. Sometimes we have so called secondary emotions, which refers to an emotional reaction to an emotion. We learn these. Some examples of these are: Feeling shame when you get angry, feeling angry when you have a shame response (for example, hurt feelings), feeling fear when you get angry (maybe you've been punished for anger). There are many more. These are *not* wired into our bodies and brains, but are learned from our families, our culture, and others. These emotions show when our needs are or are not fulfilled.

³² An approach taken into consideration by psychotherapy and as part of "awareness training" today.

³³ As primary, non-specific bodily states that can be transferred into our self-concept via a symbolization process like the one described by Rogers.

Needs are p.e.:

- Security
- Confidence
- Choice
- Faith
- Legacy
- Making a difference
- Sense of belonging/love
- Stimulus
- Influencing ... being able to influence matters in a positive way brings empowerment
- Variety
- Spirituality

Carl Rogers, Rosenberg's protégé, stresses that primary emotions (pleasure, fear, anger, sorrow and disgust – all accompanied by the same gestures and expressions across cultures and thus universally recognizable) are perceived as secondary emotions³⁴ or "pseudo-feelings"³⁵ when they cannot be assigned a place within one's self-concept or integrated into one's own self-image. If this is indeed the case, we can concentrate on the modification of primary emotions and leave aside the secondary emotions and pseudo-feelings, as these are misinterpreted affects.

2. We could attempt to manipulate primary emotions directly, on the level of neurotransmitters.

Emotions can override the ability to perceive needs that are crucial to survival. As far back as the 1950s, Olds and Milner³⁶ were able to show in an experiment with rats that, when given the opportunity to activate a switch that would release dopamine into their own bodies via an electrode implanted in their *nucleus accumbens*, the rats would activate the switch until they died of hunger or thirst. The same area of the brain has since been identified as responsible in part for how humans expe-

³⁴ Shame, frustration, irritation, curiosity, surprise, etc.

³⁵ Feeling attacked, injured, deserted, marginalized, neglected, etc.

³⁶ Olds/Milner (1954).

rience emotions like pleasure, lust and motivation. More recent studies, including Tanja Singer's, have shown that in most men, increased amounts of dopamine³⁷ are also released by the *nucleus accumbens* when they experience malicious joy (the German *Schadenfreude*), i.e. when they witness a rival who has previously acted "unfairly" towards the test person experiencing pain or loss³⁸. Oxytocin, the so-called "love hormone" seems to ameliorate this effect³⁹, which may also explain why women in the control group did not experience *Schadenfreude*, but rather empathy with their rivals. Nonetheless, *Schadenfreude* – the anticipation of which may in my opinion initiate aggressive behavior – is a reaction to a perceived (subjectively experienced) injustice.

These studies seem to indicate that a "continual state of happiness" for all people is not something we should attempt to strive for. For one, because the *nucleus accumbens* only releases increased amounts of dopamine when something is "better than expected" or when something perceived as pleasurable is also new or novel in some way⁴⁰ – thus increasing tolerance for such things and subjecting them to the law of diminishing returns. But also because (artificial) stimulation could somehow override or mask our awareness of the needs that are essential to our survival⁴¹.

If "happiness pills" and brain pacemakers⁴² aren't the answer, might it instead be sufficient to permanently eliminate the primary and negatively connoted feelings of fear, anger and sorrow along with the aggressive behavior they lead to? Assuming some sort of neuro-enhancement were possible, how could we make use of it without simultaneously endangering our own survival⁴³? And

³⁷ A phenomenon associated with experiencing pleasure or positive feelings.

³⁸ In the experiment as classically conducted, two competitors are each given the opportunity to act either fairly or unfairly. The amount of *Schadenfreude* experienced by one competitor correlates directly to unfairness previously experienced. Cf: Archival Report [<http://sans.haifa.ac.il/publications/envyschadenfreude.pdf>, retrieved 14.02.2021].

³⁹ Cf. Shamay et al. (2009).

⁴⁰ The same process which is being discussed as the basis for motivation when learning new things, cf. Spitzer (2008).

⁴¹ The same dilemma is apparent among those addicted to drugs (amphetamines, heroin, alcohol, etc.).

⁴² Such as those used to treat those suffering from Parkinson's Disease as well as depression and other neurotransmitter disorders.

⁴³ See Hans-Christian Pape's succinct remarks in the *Deutsches Ärzteblatt* in 2007: "Fear enables us to survive."

how would such technological innovation even be possible in the face of absurd applications of the postulate of refraining from aggressive behavior (see my definition of aggression as willfully harming or knowingly accepting that harm will be done to the well-being of another living creature, regardless of intent or motive)?

As for discussions of decoupling the fear network located in the limbic system and the prefrontal cortex and activating the frontal control function, those are best left for others^{44, 45}.

Do we wish, then, to accept the challenges associated with no longer viewing mankind as natural beings, but as artificial ones (works of art) which we are free to design according to certain ethic guidelines?

Regardless of how we decide, we will always be intervening in nature. Nature, as we experience it today, is concentrated art. We manipulate nature even by simply letting things happen. Should we choose not to accept that passivity itself constitutes an intervention into nature (we can't leave nature, we move about always within it, we are part of it; nature does not exist outside of our definition but is, ultimately, a manmade work of art), then our problem is solved.

But if we assume that, as natural beings (i.e. works of art) ourselves, we bear some sort of responsibility for our habitat as well as for ourselves, then we also have the freedom to determine the direction we'd like to evolve in. Do we wish to intervene in that imaginary "human nature," to reach for the tools of moral enhancement, whatever they may be, in order to dampen our primary emotions so that our empathy may hold sway? Might it not be better to simply turn off, either temporarily or permanently, those emotions with cognitive enhancement or by tinkering with our limbo-cortical networks?

Or would we rather concentrate on shaping our habitat⁴⁶ so that resources are distributed evenly, however one might define this, and the needs of all living beings can be met? So that, by giving ourselves more flexibility in choosing courses of action, the need for aggressive behavior becomes superfluous?

⁴⁴ The amygdala appears to be responsible for primitive, simple forms of associatively learned fear, while the hippocampi are responsible for more complex forms and for the context of fear. The prefrontal cortex exerts control over fear reactions.

⁴⁵ I am assuming that this is or will in the near future be within the realm of the technically possible, though whether or not we'll consider it a good thing remains open.

⁴⁶ Not only our "natural" habitat but also our social habitat.

Bibliography

- Bálcac, Honoré (2002): *Verlorene Illusionen (Illusions perdues)*. Übersetzt von Hedwig Lachmann, Leipzig
- Buber, Martin (*1979): *Das dialogische Prinzip*, Heidelberg
- Cissna, Kenneth N. (2002), *Moments of Meeting: Buber, Rogers, and the Potential for Public Dialogue*, New York
- Beck, Simon, R. (2005): *Developing nonviolent communication: an integral approach*, master thesis, University of Victoria [Victoria, British Columbia/Kanada]
- Debiec, Jacek (2005): Peptides of love and fear: Vasopressin and oxytocin modulate the integration of information in the amygdala, in: *Bioessays* (27), pp. 869–873
- Fliessbach, Klaus et al. (2007): Social comparison affects reward-related brain activity in the human ventral striatum, in: *Science* 318, pp. 1305–1308
- Kirsch, Peter et al. (2005): Oxytocin modulates neural circuitry for social cognition and fear in humans, in: *Journal of Neuroscience* 25, pp. 11489–11493
- Kurzweil, Ray (2000): *The age of spiritual machines: When computers exceed human intelligence*, London
- Lipps, Theodor (2009): *Grundzüge der Logik*, Charleston
- Olds, James/Milner, Peter (1954): Positive reinforcement produced by electrical stimulation of the septal area and other regions of rat brain, in: *Journal of Comparative and Physiological Psychology* 47, pp. 419–427
- Rogers, Carl (1994): *Person to Person: The Problem of Being Human*, London
- (1995): *On Becoming a Person: A Therapist's View of Psychotherapy*, New York
- Rogers, Carl/Skinner, Burrhus (2011): *Some Issues Concerning The Control of Human Behavior*, kindle edition, www.all-about-psychology.com
- Rosenberg, Marshall B. (2005): *Speak Peace in a World of Conflict: What You Say Next Will Change Your World*, Encinitas
- Sánchez Bodas, Ricardo Andres (2013): *Manifiesto humanístico*, kindle edition, edicioens LEA
- Schneider, Kurt (1923): *Die psychopathischen Persönlichkeiten*, Wien (Nachdruck: Leipzig 1950)
- Saß, Henning/Herpertz, Sabine, Hg. (1999): *Psychotherapie der Persönlichkeitsstörungen*, Stuttgart–New York
- Shamay-Tsoory, Simone G. et al. (2009): Intranasal administration of oxytocin increases envy and schadenfreude (gloating), in: *Journal of Biological Psychiatry* 66, pp. 864–870
- Siep, Ludwig (1999): *Ethik und Menschenbild*, Münster
- Soentgen, Jens (2014): Der Lech als Cyborg, in: Marita Kraus u. a., Hg., *Der gezähmte Lech. Fluss der Extreme*, München, S. 151–160

Spitzer, Manfred (2008): *Selbstbestimmen: Gehirnforschung und die Frage: Was sollen wir tun?*, Heidelberg

SymbioticA (2013): *Semipermeable+*, UWA, Perth

Online Sources, retrieved 14.02.2021

Robert Gonzales, NVC Training Institute; www.nvctraininginsitute.com, 2004

http://www.cnvc.org/sites/cnvc.org/files/NVC_Research_Files/NVC%20Research/Beck-Developing_NVC_Integral_Approach.pdf

<http://www.aerzteblatt.de/archiv/55491>

<http://examples.yourdictionary.com/ethical-dilemma-examples.html>

<http://sans.haifa.ac.il/publications/envyschadenfreude.pdf>

<http://www.spektrum.de/alias/wahrnehmungspsychologie/die-schoenheitsformel/793179>

<http://www.symbiotica.uwa.edu.au/>

<http://www.who.int/classifications/icd/en/GRNBOOK.pdf>

Warum Data malt – Interdisziplinarität und Ästhetik

Stefanie Voigt

Zusammenfassung

Das Rätsel des menschlichen Bewusstseins kann im Rahmen einer interdisziplinären Ästhetiktheorie erfolgreich bearbeitet werden. Zwar ist diese Disziplin aus geschichtlichen Gründen derzeit noch randständig, kann aber in einer modernen, systemtheoretischen Gestalt zeigen, was menschliches Bewusstsein mit drei Stufen von Schönheitsempfinden zu tun hat: einfacher Wiedererkennung von Mustern, intensiver Mustersuche und Ekstase oder Enstase.

Dass eine intuitive Ahnung diesen Zusammenhang von Ästhetik und Bewusstsein nahelegt, belegt das populärkulturelle Beispiel des Androiden Data aus *Star Trek, The Next Generation*, der Kunst betreibt, um menschlich zu werden.

Abstract

The riddle of human consciousness can successfully be approached in the context of an interdisciplinary theory of aesthetics. This discipline, however, still is marginalized for historical reasons. But in a modern shape based on the theory of systems, that discipline can show how human consciousness is connected to three steps of the experience of beauty: simple recognition of patterns; intensive search for patterns; ecstasy or “enstasy.” That there is an intuitive grasp of this connection is elucidated by the example of the android Data from *Star Trek, The Next Generation*, who takes to the arts in order to become human.

1. Zur Relevanz der Ästhetik

Auf der Suche nach dem Geheimnis der menschlichen Seele wird die Ästhetik meist außer Acht gelassen. Zumindest spielt sie als ernst genommener Gesprächspartner im interdisziplinären Kanon der Kognitionswissenschaften keine Rolle. Gerade dadurch aber scheint das Bewusstsein zu dem unauflösbaren Rätsel zu werden, als welches es nach wie vor weithin gilt. Aber – so die hier vertretene These¹ – eine systemtheoretische Neuformulierung der Ästhetik sowie der an sie angrenzenden Nachbardisziplinen ermöglicht eine neue Vorstellung von Ästhetik, genauer gesagt: ein Modell der menschlichen Informationsverarbeitung, das Bewusstsein als ein ästhetisches Phänomen definiert. Dieses Modell kann aufgrund seiner Formalisierung von ästhetischer Wahrnehmung beziehungsweise von Bewusstsein damit zusammenhängende Themengebiete, die bislang als diskursiv nicht fassbar, weil ‚künstlerisch‘ galten², dem disziplinären wie interdisziplinären Dialog erschließen.

Es gibt in verschiedenen Disziplinen viele Ansätze, Ästhetik als koordinierende Kerndisziplin der Bewusstseinswissenschaften zu etablieren³, und diese Ansätze weisen erstaunliche inhaltliche Konvergenzen auf – ein Indiz für das noch weiterhin unausgeschöpfte interdisziplinäre Potenzial der Ästhetik⁴. Trotzdem bleibt, während zwar zahlreiche Disziplinen interdisziplinäre Familienzusammenführungen feiern, die Ästhetik weitgehend außen vor wie ein ungeliebtes Kind⁵. Dabei gab und gibt es bis heute namhafte Sprecher, die die sinnliche Wahrnehmung

¹ Die Thesen dieses Beitrags sind ausführlicher dargelegt in Verf. (2005). Dort finden sich auch zahlreiche, in den folgenden Fußnoten aufgeführte Belegstellen, deren neuerliche Wiedergabe den Rahmen dieses Beitrags sprengen würde.

² Vgl. ebd., S. 106f.

³ Vgl. ebd., S. 25f.

⁴ Beispielsweise sieht die Kunsttheoretikerin Brandstätter (2008, S. 65–67) „Transdisziplinarität als Chance“ und versteht darunter auch interdisziplinäre Erforschung ästhetischer Phänomene. Auch Franz (2004, S. 133) verlangt nach einer „Ästhetik [...], die sich im Spannungsfeld von Philosophie, Einzelwissenschaften und Technik-Diskursen ansiedelt und zu behaupten vermag“.

⁵ So grenzt beispielsweise Reicher (2005, S. 24f.) die ästhetische von der empirischen und damit auch der psychologischen Ästhetik scharf ab, fragt aber nicht nach Perspektiven interdisziplinärer Ergänzung. Krohn (2006) ist zwar unter dem Stichwort „Ästhetik in der Wissenschaft“ dem „interdisziplinäre[n] Diskurs über das Gestalten und Darstellen von Wissen“ gewidmet (so Titel und Untertitel), belässt es jedoch weitgehend bei einem Nebeneinander einzelner Disziplinen. Ein transkribiertes Gespräch verschiedener Fach-

und das Empfinden von Schönheit als zentral für das menschliche Bewusstsein erachten⁶. Diese Praktiker und Theoretiker der Ästhetik definieren ihren Gegenstand als grundlegende Technik der menschlichen Informationsverarbeitung, die weit mehr als repräsentativen Wandschmuck betrifft. Vielmehr ist ihnen zufolge hier so etwas verborgen wie die Weltformel aller Geisteswissenschaften, das Geheimnis um die menschliche Seele, um Glück, Schönheit und Lebendigkeit – also die Klärung all der Begriffe, die ihre Heimat in wissenschaftlichen – insbesondere naturwissenschaftlichen – Diskursen unter dem Einfluss des Positivismus verloren hatten⁷. Allerdings konnten jene Autoren diesen Trend nicht wenden. Daher wurde die Ästhetik von bewusstseinstheoretischen Diskursen bislang eher vernachlässigt und erweckte darüber hinaus den Eindruck manchmal fast solipsistischer Abgeschlossenheit. Für diese Situation lassen sich die folgenden acht Gründe anführen⁸.

2. Gründe für die Randstellung der Ästhetik

Kunst als der hauptsächliche Gegenstand der Ästhetik wird insbesondere seit der Romantik häufig dadurch definiert, dass sie nicht wissenschaftlich, sondern auf eine andere, „intuitive“, Weise erfassbar ist und sich somit auch nicht interdisziplinär in einen wissenschaftlichen Fächerkanon integrieren lässt (1. Grund)⁹. Zudem gilt Kunst nach einer weit verbreiteten philosophischen Begriffsbestimmung als zweckfrei *par excellence*¹⁰ (2. Grund). Diese begriffliche Festlegung

vertreter endet aber mit der versöhnlichen Feststellung, sie könnten noch viel voneinander lernen (ebd., S. 197f.).

⁶ Es sind dies – jeweils unter vielen anderen – auf dem Feld der Kunst selbst Leonardo, Cézanne, Malewitsch, Picasso und Beuys; in der Psychologie James, Jaynes, Festinger und Beyer; in der Philosophie Adorno, Lyotard, Gadamer, Sloterdijk; in Linguistik und Semiotik Chomsky, Ong, Peirce, Wittgenstein und Bachtin, in der Anthropologie Bateson, Lurija, Duerr und Harris, in der Kunstgeschichte bzw. -theorie Flusser, Barthes, Panofsky, Sedlmayr und Bataille. Vgl. Verf. (2005), S. 25f.

⁷ Vgl. Städtke (2001), S. 483.

⁸ Vgl. Verf. (2005), S. 19ff.

⁹ Vgl. ebd., S. 22, und Städtke (2003), S. 632–634.

¹⁰ Vgl. Verf. (2005), S. 104f. sowie Reicher (2005), S. 151f. Dadurch wird Kunst natürlich zum Gegenstand von Positionen, die übergeordnete Zwecke aller menschlichen Aktivitäten propagieren; vgl. Hammermeister (2007), S. 158–162.

stellt Kunst vor ein Dilemma: Genügt sie dennoch irgendeinem Zweck, wird sie als „bloßes“ Design in Form von Kitsch oder Kunsthandwerk ausgegeben¹¹. Gibt sich Kunst dagegen tatsächlich als zweckfrei, gerät sie schnell in den Verdacht einer schöngestigen Luxusbeschäftigung (3. Grund)¹². Die Kunstphilosophie vermag in diesem Konflikt nicht zu vermitteln, da sie paradoxerweise kaum direkten Kontakt zu ihrem eigenen Gegenstand, der Kunst hat. Daher widmen sich die Kunstphilosophie zunehmend den Reflexionen über ihren eigenen Status¹³. Darüber kommt selbst elementare begriffsanalytische Arbeit zu kurz, allem voran eine saubere Trennung zwischen den Begriffen der Ästhetik und der Schönheit, die auch in der Kunstphilosophie oft synonym gebraucht werden und doch häufig sehr Verschiedenes bedeuten (4. Grund)¹⁴. Denn nicht alles, was als ästhetisch tituliert wird, ist auch schön. Und nicht viel von dem, was die einen als schön bezeichnen, würden andere überhaupt als ästhetisch bezeichnen, v. a. der überholt bildungsbürgerlichen Attitüde des Begriffs Schönheit wegen (5. Grund)¹⁵. Angesichts solcher Verwirrungen schon im Hinblick auf die grundlegenden Begriffe ist es nicht verwunderlich, dass die Kunstphilosophie aus einer uneinheitlichen Gemengelage verschiedener Meinungen zu den Themen Ästhetik und Schönheit besteht (sechster Grund)¹⁶.

Wo so viel Konfusion herrscht, liegt es nahe, um einer Therapie willen die Psychologie zu konsultieren. Diese aber weist den dazu erforderlichen Dialog notgedrungen zurück. Die naturwissenschaftskonforme Psychologie unserer Zeit widmet sich lieber Gegenständen, die sich quantifizieren und statistisch erfassen lassen. Dabei kommen zwar auch Aussagen über Ästhetik zustande, diese haben aber sehr elementaren Charakter. So wird z. B. ermittelt, dass die bevorzugte Wahl der Farbkombination schwarz-gelb eher Neurosen indiziert als jede andere Farbzusammen-

¹¹ Vgl. Verf. (2005), S. 108–110.

¹² Vgl. ebd., S. 29f. Dies wird satirisch thematisiert von Kishon (1986).

¹³ Vgl. Welsch o. J., S. 1. Welsch vergleicht die Beziehung zwischen Kunst und Kunstphilosophie ebd. mit einer gescheiterten Ehe und sieht daher die beste Lösung in einer einvernehmlichen Trennung der beiden. Zur entsprechenden Entfremdungsgeschichte vgl. Franke (2007).

¹⁴ Vgl. Verf. (2005), S. 20.

¹⁵ Vgl. ebd., S. 20f.

¹⁶ Vgl. ebd., S. 21f. Reicher (2005), S. 9–31, verdeutlicht diese begrifflichen Probleme in einer begriffsanalytisch angelegten Ästhetik.

stellung¹⁷. Elementarpsychologie dieser Art ist zwar interessant (v. a. für Bienen und Feuersalamander), für große Einsichten in das Wesen des Menschen ist sie aber zu speziell. Indem Psychologie den Großteil ästhetischer Phänomene als wissenschaftlich nicht fassbar abschreibt, folgt sie damit der Mystifikation der Kunst als etwas Unaussprechlichem (7. Grund)¹⁸.

Durch den Verzicht auf „große“ ästhetische Theorien erspart sich die heutige Psychologie immerhin Angriffe seitens der Verfechter einer historischen Anthropologie wie etwa der sog. Annales-Schule um Le Goff¹⁹. Nach dieser Position wandelt sich das Wesen des Menschen im Lauf der Zeit, und daher wäre es ohnehin verfehlt, allgemeine wissenschaftliche, wie etwa psychologische, Aussagen über „die“ ästhetische Wahrnehmung etc. zu machen. Dass sich aber gerade mit Hilfe der Ästhetik die Dynamik der menschlichen Psyche durch ihre verschiedenen geschichtlichen Wandlungen hindurch erklären lässt, wird aus jener historistischen Perspektive alleine nicht zugänglich. Und dies gilt nicht nur für die genannten Disziplinen, sondern insgesamt: Ästhetik ist eben keine Disziplin neben vielen anderen, sondern ein interdisziplinärer Forschungsbereich. Gerade daher haben die einzelnen Disziplinen, die zunächst auf ihr jeweiliges Gebiet verwiesen sind, Schwierigkeiten mit dem Zugang zur Ästhetik (8. Grund)²⁰.

3. Das Menschliche

Ästhetik ist vielen Autoren vieler Disziplinen zufolge der Schlüssel zum Bewusstsein, die Verständigung über die Grenzen der Disziplinen hinaus ist aber schwer. Um in dieser scheinbar verfahrenen Situation weiterzukommen, liegt es, wie so oft bei voneinander abgegrenzten Systemen nahe, einen völlig Außenstehenden nach seiner Einschätzung zu fragen. Als gutes Anschauungsobjekt eignet sich in dieser Hinsicht der malende und geigende Androide Data aus der Fernsehserie

¹⁷ Vgl. ebd., S. 23.

¹⁸ Vgl. ebd., S. 228.

¹⁹ Vgl. ebd., S. 46.

²⁰ Vgl. ebd., S. 23f.

„Star Trek – The Next Generation“²¹. Er entwickelt immer dann menschliche Eigenschaften wie ein Gewissen und Individualität, wenn er sich mit Kunst beschäftigt. Erst dann gerät für ihn auch das Lachen oder eine bestimmte Art von genießerischer Selbstgenügsamkeit in Reichweite, die sonst nur Menschen zuzukommen scheint. Allerdings wird Data dadurch auch anfällig für Selbstzweifel oder Angst vor sprichwörtlich „menschlichem“ Versagen. Hier wird offenbar eine weit verbreitete, wenn auch bislang kaum explizierte Überzeugung in Szene gesetzt: Menschen kennen Emotionalität, Mitleid, Schuld, Einfühlungsvermögen, Reue, Freude und Trauer oder Schmerzfähigkeit, Ironie und Kreativität, Selbstverantwortlichkeit. Diese und weitere ‚typisch menschliche‘ Phänomene sind eng mit dem Bereich der Ästhetik verbunden, werden zumindest in ihm auf eine exemplarische, intensive Weise erlebbar. Insbesondere gibt es in diesem Bereich nicht das so genannte „computertechnische Paradoxon“²²: Die bisher konstruierten Datenverarbeitungsmaschinen stürzen schlichtweg ab, wenn sie mit widersprüchlichen Informationen konfrontiert werden, die in ihrer Programmierung nicht vorgesehen sind. Ästhetik dagegen lebt geradezu von derartigen Widersprüchen, die in ihr im Gegenteil auf vielfältigsten Ebenen zur Geltung gebracht und verstärkt werden²³. Wenn sich ein künstliches Gebilde mit Ästhetik befasst, sind wir daher dazu geneigt, ihm auch mehr oder weniger große Menschlichkeit zuzuschreiben. Nach der Meinung einiger Psychologen zeichnet sich Bewusstsein sogar gerade durch den kreativen Umgang mit Widersprüchen aus²⁴. Demnach ‚funktionieren‘ Menschen, solange sie am Leben sind, und sie haben das Gefühl zu leben, solange sie Dinge als schön empfinden. Denn sobald ein Mensch, aus welchen Gründen auch immer, nicht mehr dazu in der Lage ist, etwas schön zu finden, wird er krank²⁵. Auch dies unterscheidet Menschen von Computern, und auch das spricht für einen Zusammenhang zwischen menschlichem Bewusstsein und Schönheitsempfinden. Die Krankheitsbilder von Suizidalität, Schizophrenie und Epilepsie wurden bisher weitestgehend nur neurophysiologisch vermessen,

²¹ Zu Data als einem zwar hypothetischen, nichtsdestoweniger aber aufschlussreichen Testfall für philosophische Fragestellungen vgl. Alexy (2000), S. 1.

²² Vgl. Draisma (1996).

²³ Vgl. Verf. (2005), S. 119–122.

²⁴ Vgl. Jaynes (1988); Dörner (1999).

²⁵ Vgl. Verf. (2005), S. 147–149. Daher versteht Nietzsche die Kunst als „Stimulanz zum Leben“; zit. nach Peitsch (2001), S. 193, Anm. 119.

ihre inhaltliche Logik bzw. die geistige Regulierung innerer Zustände wurde hingegen noch nicht umfassend untersucht. In jedem dieser Fälle vermögen die Betroffenen ihre Vorstellung von der Welt, das innere Bild ihrer Umwelt, also das Produkt ästhetischer Wahrnehmung, nicht mehr auf konventionelle Art zu regulieren. Dann praktizieren sie zwangsläufige Alternativ-Manöver, nehmen schlimmstenfalls Sachen wahr, die gar nicht da sind oder produzieren spontan extreme Glücks- oder Schönheitsgefühle, die das psychische System unter Umständen auch überfordern, wenn nicht gar zur Selbsttötung motivieren. Würden die bislang vorgelegten Bewusstseinssimulationen wirkliches Bewusstsein beinhalten, könnten sie von Glück reden, dass ihnen diese Probleme erspart blieben, allein, dann würden sie auch darüber trauern, keine Kunst und keine Literatur zu haben. Denn diese Bereiche strotzen vor solchen Problemen, vor Emotionen und Schönheiten verschiedenster Couleur. Die Informationsverarbeitungssysteme der Menschen scheinen größer zu sein als die Summe ihrer Einzelteile. Aber wie ist das möglich? Dies bleibt bis her ungeklärt. Neurophysiologen vermessen das Gehirn und Informatiker programmieren ihre Computer und woanders disputieren die Geisteswissenschaftler über die menschliche Seele und die hehre Kunst – und dazwischen klafft ein Abgrund, den noch kein Netzwerk überbrücken konnte, ungeachtet vielfältiger Versuche und Zusammenschlüsse von Informatikern mit Neurologen, Ethikern, Bildwissenschaftlern oder Neurobiologen zur Schaffung neuer Disziplinen, wie z. B. Neuroinformatik oder Neuroethik.

4. Ein Modell des Menschlichen

Was hat Data wohl getan, um das Geheimnis des Menschlichen in der Kunst zu ergründen? Hat er eine Einführung in die Ästhetiktheorie von Platon bis Bazon Brock eingescannt und in die Sprache seiner künstlichen Synapsen übersetzt? Da es sich beim Gehirn Datas um einen Digitalcomputer handelt, würde das Ganze mit basalen Mitteln der digitalen Informationsverarbeitung beschrieben, mit dem Ziel, alle vorliegenden partiellen disziplinären Erkenntnisse im Rahmen einer einzigen Theorie zu vereinigen. Eine derartige Übersetzung hehrer Kunst-Theorien in den digitalen 0-1-Code ist trotz des genannten Kontrastes zwischen Computern und Menschen vielleicht tatsächlich möglich. Die Vielzahl individueller und epochegebundener, häufig widersprüchlicher Ästhetiktheorien kann im Rahmen

einer systemtheoretisch orientierten Psychologie, die sich auf die Einsichten der Anthropologie stützt, zumindest auf acht einheitliche Rahmen-Variablen reduziert werden. Unter Verwendung des aristotelisch-wienerschen-batesonischen Begriffs des Unterschieds²⁶ (als kleinster Einheit jeglicher geistigen Leistung) lässt sich ästhetisches Empfinden folgendermaßen charakterisieren:

1. In ästhetischem Empfinden geschieht eine Vermittlung zwischen äußeren Eindrücken und inneren Interpretationsmustern, einem Zusammentreffen von sinnlicher Wahrnehmung und abstraktem Denken bzw. gleichzeitige Wahrnehmung von Welt und eigener Person, von subjektiv inneren Schemata und subjektiv äußerer Welt – egal ob Platon zwischen Ideen und Erscheinungen differenziert oder Aristoteles zwischen einem Kunstwerk und dem sich damit identifizierenden Betrachter.
2. Diese allgemeine Vermittlung funktioniert im Detail über den Vergleich innerer und äußerer Muster, einer Entsprechung und Nachmodellierung von Strukturprinzipien – und zwar bei Ficinos Spiegelprojektion genauso wie bei Albertis Proportionsstudien.
3. Was durch die interpretierende Wahrnehmung dieses Vergleichs, also einer Wahrnehmung der Wahrnehmung entsteht, sind lustbringende Erlebnisqualitäten, der immer wieder beschworene Eigenwert und Erkenntnisgehalt des Ästhetischen – wie Edmund Burkes „pleasure“ oder Kants „interesseloses Wohlgefallen“.
4. Diese Gefühlsqualitäten zeigen sich in verschiedenen starken Ausprägungen, von Interesse oder Faszination bis hin zu Ekstase²⁷ oder Enstase – von Lessings „Rührung“ bis hin zu Stendhals „Syndrom“.
5. Möglich ist das alles stets nur durch eine gemäßigt stressfreie, weder als langweilig, noch als zu aufregend empfundene Art der Wahrnehmung – bei Schopenhauers „Naturbetrachtung“ genauso wie bei Nietzsches Überlegenheitsgefühl des „Übermenschen“, oder der modernen Abject Art²⁸.
6. Ein Bereich eigener Regelhaftigkeit entsteht, da die Rezeption aufgrund dieser kontemplativen Art der Rezeption nicht bewusst und rational dis-

²⁶ Vgl. dazu und zum folgenden Verf. (2005), S. 91ff.

²⁷ Vgl. Kaden (2002), S. 263.

²⁸ Vgl. Menninghaus (2001), S. 175f.

tanziert erfasst werden kann und darf – darum spricht Boileau vom „je ne sais quoi“ („ich weiß nicht was“) und Goodman von den „eigenen Sprachen der Kunst“.

7. Dieser Vorgang garantiert in seiner Gesamtheit das langfristige Funktionieren des Systems Psyche durch bessere Denkfähigkeit nach außen und durch emotionalen Lustgewinn nach innen – darum ist Lyotards „Präsenz“ ebenso wichtig wie Wittgensteins „richtige Perspektive“.
8. Die Motivation für ästhetische Betrachtung oder sogenanntes diskursives Denken hängt ab von situativer und persönlicher Vernetzungsfähigkeit – von Baumgartens „Anlage zur ästhetiklogischen Erkenntnis“ oder Schillers „Spieltrieb“.

Es ist möglich, diese acht Aussagen in Gestalt eines Flussdiagramms zu formalisieren und dadurch zu einer funktionalen Beschreibung der entsprechenden psychischen Prozesse zu gelangen, die ästhetische Wahrnehmung wertneutral als Wahrnehmung zwischen Bestimmtheit und Unbestimmtheit definiert²⁹. Bestimmtheit beruht auf der Menge bereits systemimmanent vorhandener Muster; Unbestimmtheit ist alles, was diesen Mustern nicht entspricht. Zwischen diesen beiden Instanzen herrscht zunächst ein kategorialer Gegensatz – das Bestimmte ist nicht das Unbestimmte und vice versa. Dieser Gegensatz lässt sich jedoch überbrücken, indem eine Vermittlung zwischen Bestimmtheit und Unbestimmtheit stattfindet, indem das System seine bisherigen Muster umstrukturiert und solchermaßen neu interpretiert, oder aber, falls dies nicht ausreicht, neue Muster schafft. Dieser Prozess kann sich in verschiedenen Zyklen und über zunehmend große Bereiche des gegebenen Mustervorrats hinweg wiederholen, bis er entweder zu einem Erfolg führt oder das daran beteiligte System überlastet, was zum Abbruch der Suche führt. Dieses zunächst sehr formale und abstrakte Modell kann als Deutung von Phänomenen dienen, die von verschiedenen Disziplinen bisher auf unterschiedliche Weise beschrieben wurden, und gerade so die zum Verständnis des Bewusstseins erforderliche interdisziplinäre Integration leisten:

- *Neurophysiologisch* können anhand verschiedener Regulationen des Botenstoffs Dopamin zwei verschiedene, komplementäre Arten der Informations-

²⁹ Vgl. Verf. (2005), S. 188 u. ö.

verarbeitung nachgewiesen werden, nämlich sogenanntes fixatives bzw. vagatives Denken. Fixatives Denken operiert mit abstrakten, vereinfachenden Begriffen und äußert sich dementsprechend auch in einer Eindämmung zerebraler Aktivität. Vagatives Denken dagegen verfährt sprunghaft-vielschichtig, gleichsam irrational, und äußert sich in einer zunehmenden Ausweitung der Gehirntätigkeit. Die EEG-Messungen des Letzteren zeigen Parallelen zwischen Meditationstechniken oder anderen kulturellen Formen intuitiv-ästhetischer Praxis, sei es bei amerikanischen Konzertbesuchern oder sibirischen Schamanen.

- *Anthropologisch* gesehen zählt die Fähigkeit zu solchen Ekstasen in den meisten Kulturen zu den Grundbedingungen des gesunden Menschenverstandes. Denn nur zeitweilige Ekstasen reinigen das Denken so, dass es sich immer wieder neuen Anforderungen stellen kann, ohne sich dogmatisch zu verengen.
- *Mentalitätsgeschichtliche* Vergleiche zeigen allerdings, dass die Ekstasefähigkeit im Lauf der Moderne zunehmend verinnerlicht und säkularisiert wurde, was sie in einen Sonderbereich des Ästhetischen, die gesellschaftlich anerkannte Kunst, abgedrängt hat.
- *Kunst, Kunstgeschichte und Literaturwissenschaften* bieten zahlreiche einschlägige Beispiele dafür, etwa das zeitgleiche Auftreten der Verinnerlichung des Erlebens und gesellschaftlicher Individualisierungsprozesse, die Parallele von Todestabuisierung und gleichzeitiger Dramatisierung des Todes, die Verwissenschaftlichung des Denkens während der Entdeckung des Sujets der atmosphärischen Stimmung o. ä.

So entsteht ein Modell, das Schönheitsempfinden auf drei Stufen erklärt: einfache Schönheit in Form reiner Wiedererkennung von Mustern (das meist auch als bloßer Kitsch interpretiert bzw. von vielen als langweilig empfunden wird), eine zweite und nur mit kognitiven Schwierigkeiten zugängliche Ebene intensiver Muster-suche (die Rezeption wird dann als „faszinierend“ oder „interessant“ beschrieben, während Neuinterpretationen des kognitiven Problems erarbeitet werden – oder die ästhetische Mustersuche wird abgebrochen), und einer dritten Stufe der Ekstase oder Enstase. Die Nähe zu diesem Bereich wird indiziert anhand Lessingscher „Rührung“ bzw. physiologischer Tränen. Was auf dieser Stufe selber passiert, wurde historisch gerne als göttlich beschrieben, in der Neuzeit wird daraus die Erfahrung von Wahrheit bzw. totaler Schönheit. Hier verschwimmen die Grenzen zwischen

Ich und Welt genauso wie andere Formen von Unterschieden. Der „Rückweg“ zum Normalbewusstsein wird traditionell gerne als „Auferstehung“ tituliert³⁰; misslingt dieser Rückweg, kommt es zur Schizophrenie, in der man den Wald vor lauter Reizen nicht mehr sieht.

Auf diese Weise können alle wesentlichen Topoi der Ästhetik funktional erklärt werden, angefangen bei Max Ernsts „Mut des Künstlers“ bis hin zum Klischee des genialen Künstlers als Opfer seiner Triebe. Weitere Beispiele sind die Metaphern von Kindlichkeit, Rausch, Traum und Sex, die Schönheit von Idealisierungen wie auch die Schönheit der Hässlichkeit, ebenso wie die verschiedenen Arten der Empathie, sei es Empathie gegenüber Menschen, ozeanischen Weiten oder luxuriösen Automobilen. Melancholie, mystische Reizüberschwemmungen und lustvolle Selbstausschönungen, die Legende von der Zweckfreiheit und Unbeschreibbarkeit der Kunst: All dies funktioniert über die Logik der Unterschiedsregulierung im menschlichen Denken, einschließlich der Unterschiede zwischen den Gegenständen des Denkens, zwischen dem Menschen und seiner Umwelt und zwischen dem Denken in Unterschieden und dem unterschiedslosen Denken ästhetischer Grenzerfahrungen. Dieser letzte Unterschied kann nicht anders, als zweckfrei und unbeschreibbar sein, sonst wäre er nicht unterschiedslos, sondern so fixativ und rational zu beschreiben, wie alles andere auch. Schönheit spottet jeder Beschreibung, und dafür hat sie gute Gründe. Denn Ästhetik funktioniert wie der blinde Fleck der Wahrnehmung. Blind, aber notwendig, damit das Auge funktioniert. Denn der Lustgewinn durch den „Kurzschluss“ des Denkens in wenigen unterschiedslosen Momenten schafft einen notwendigen und bewussteinserzeugenden Gegenpol zum ‚Normalbewusstsein‘, der diesem zugleich auch neue Motivation verleiht; denn jede Abstraktion setzt das Wissen um das Gegenteil voraus. Jedes Pferd ist definiert durch alles, was Nicht-Pferd ist, und jedes klare Denken durch sein Gegenteil. für das wirkliche Leben: Ohne Ästhetisches wird der Alltag öde, umgekehrt ist nur Ästhetisches eine Steilvorlage für den Wahnsinn. bei Kant: „Ohne Sinnlichkeit würde uns kein Gegenstand gegeben und ohne Verstand keiner gedacht werden. Gedanken ohne Inhalt sind leer, Anschauungen ohne Begriffe sind blind“³¹. Die Systemtheorie postuliert zwar, dass sich kein System selber erkennen kann, weil es immer Teil desselben ist. Aber die Systemtheorie der philosophischen

³⁰ Vgl. ebd., S. 192, sowie Funke (1996), S. 121ff.

³¹ Kant: Kritik der reinen Vernunft, B 75/A 51.

Ästhetik beschreibt Ekstase, die Extremform ästhetischen Empfindens, an vielen Stellen als Möglichkeit zur Selbstbegegnung, der sog. Heautoskopie³². Kein Argument ist gegen Zweifler gefeit, doch in jedem Fall schaffen solche Erfahrungen zumindest potenzielle Möglichkeiten zur Selbstdistanz durch Annäherung. Maturana und Varela zufolge ist der Gebrauch dieser Chance auf Selbstdistanzierung ein Indiz und eine Voraussetzung für Intelligenz³³. Hat die Evolution also das Schönheitsempfinden geschaffen, um Intelligenz zu ermöglichen? Unwahrscheinlich ist das nicht, nur ist es nicht die landläufig so genannte Art von Schönheit, die die Menschheit weiterbringt. Einstein hat einer verbreiteten Anekdote nach nicht umsonst den Heiratsantrag eines Models abgelehnt. Er argumentierte mit der Möglichkeit gemeinsamer Kinder mit der Intelligenz der Mutter und dem Aussehen des Vaters. Doch auch die Forschungen zu Menschen, die nach gesellschaftlich verbreiteten Maßstäben als schön gelten, zeigen, dass Aussehen nicht alles ist³⁴. Im Gegenteil werden attraktive Menschen laut Statistik im Leben viel wahrscheinlicher unglücklich als Durchschnittstypen. Im beschriebenen Modell der ästhetischen Wahrnehmung spielen darum nicht nur der formale, sondern auch der inhaltliche Aspekt eine große Rolle, und es zeigt sich deutlich, dass der Mensch weder reines Denken noch reine Sinnlichkeit ist, sondern eben eine eigendynamische Mischung aus beidem. In Zeiten der Industrialisierung des sogenannten „Menschenpotentials“ zeigt ein interdisziplinäres Ästhetikmodell, dass die Vernachlässigung der musischen Fächer nicht unbedingt von Vorteil für die geförderten anderen Fächer ist. Nicht nur die Kontemplativität des Musischen, sondern auch Fairness und Aufrichtigkeit zeigen sich als wichtig für den Zugang zum Ästhetischen. Dass dieses Modell formal wie eine Bauanleitung für Künstliche Intelligenz aufbereitet wurde, heißt nach ihm immanenten Logik aber nicht, dass diese Anleitung auch wirklich in die Tat umgesetzt werden sollte, so wie auch jeder Versuch einer endgültigen Definition von Ästhetik ein grobes Scheitern am Thema wäre – weil es den obligatorischen Anteil an rational nicht zugänglicher Unbeschreibbarkeit zerstören würde. Auch würden bei einem korrekten Nachbau Zufälle und Mutationen das Ergebnis so unberechenbar machen wie den Menschen selber, nur dann wäre der Nachbau getreu. Gibt es im Gehirn nämlich wirklich so etwas wie eine „chaotische

³² Vgl. Verf. (2005), S. 205.

³³ Ebd., S. 139.

³⁴ Vgl. Menninghaus (2003).

Kausalität“³⁵, dann wäre jede Vorhersage von was auch immer völlig unmöglich. Daraus ergäbe sich eine unableitbar große Menge an Möglichkeiten von Erlebnissen und internen Vernetzungsmöglichkeiten und damit eine entsprechende Menge an Vergessensmöglichkeiten. Spätestens an dieser Stelle käme die Frage nach dem Sinn des Unterfangens auf. Schließlich ist der eigene Nachbar schon unverständlich und intransparent genug und noch dazu bereits vorhanden. Also ist Ästhetisches nicht nur aus programmatischen, sondern auch aus pragmatischen Gründen kaum programmierbar – und mit dieser Überlegung reichen sich Naturwissenschaft und Kunst die Hände. Trotzdem bietet ein solches Modell eine Ausgangsbasis für genauere Definitionen in verschiedenen Disziplinen und Anreiz für weitere Forschungen³⁶. Beispielsweise zitiert der psychologische Begriff des „Möglichkeitssinns“ im Grunde genommen den Begriff der „Möglichkeiten“ bei Dionysios Areopagita, der reale Dinge immer nur als „Möglichkeiten“ des Schönen und Guten beschreibt. Ebenso nimmt Plotins unterschiedsloses Eines als Ort der religiösen Erfahrung bestimmte neurophysiologische Ergebnisse vorweg, nämlich die bei bestimmten psychologischen Abläufen eintretende Verringerung der Hirnaktivität für die Unterscheidung zwischen Selbst und äußerer Welt. Durch die Analyse solcher Strukturähnlichkeiten ließe sich der bisherige, sehr breite Ästhetik-Begriff präzisieren. Dies betrifft auch bestimmte sozial-ethische Bewertungen: Aus vorneuzeitlicher Sicht gilt z. B. ein moderner Mensch mit geringem Gesellschaftsbezug als verrückt. Der historische Mensch aus der Sicht seines modernen Nachfolgers allerdings auch, schon alleine seiner ekstatischen Praktiken oder ‚verkehrter Welten‘ wegen, die gerade in traditionellen Kulturen als Umkehr der überlieferten Ordnung selbige stabilisieren eine wichtige Rolle spielen (etwa die Saturnalien oder der Karneval). Auch das Schlagwort der Überästhetisierung setzt einen klar definierten Begriff des Menschen samt bestimmter Wertungen voraus, die bei Vorlage eines solchen Modells eventuell eher wissenschaftlich formuliert werden können als ohne – wenngleich das Modell eigentlich schon seinen künstlerischen wie auch systemtheoretischen Ansprüchen Genüge täte, wenn nur die Vorstellung solcher Interdisziplinarität Spaß am Denken in Aussicht stellte, denn es ist nicht auszuschließen, dass Data seine Recherche auf diese Grundlage gestellt hätte.

³⁵ Vgl. Maturana/Varela (1987).

³⁶ Zum Folgenden vgl. ebd., S. 205ff.

Literatur

- Alexy, Robert (2000): *Data und die Menschenrechte. Positronisches Gehirn und doppeltriadischer Personenbegriff*. (Vortrag, gehalten am 08.02.2000 im Auditorium Maximum der Christian-Albrechts-Universität zu Kiel im Rahmen einer Ringvorlesung „Populärkultur am Beispiel von Star Trek“) [<https://www.alexey.jura.uni-kiel.de/de/download/data-und-die-menschenrechte>, Zugriff: 01.02.2021]
- Brandstätter, Ursula (2008): *Grundfragen der Ästhetik. Bild – Musik – Sprache – Körper*, Weimar–Wien
- Dörner, Dietrich (1999): *Bauplan für eine Seele*, Hamburg
- Draisma, Douwe (1996): *Die Metaphernmaschine. Eine Geschichte des Gedächtnisses*, Darmstadt
- Franke, Ursula (2007): Nach Hegel – Zur Differenz von Ästhetik und Kunstwissenschaft(en), in: Josef Früchtel/Maria Moog-Grünewald, Hg., *Ästhetik in metaphysikkritischen Zeiten. 100 Jahre „Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft“*, Hamburg, S. 73–91
- Franz, Michael (2004): Ästhetik zwischen Philosophie, Wissenschaftsdisziplin und Technologie-Diskursen. In: Karin Hirdina/Renate Reschke, Hg., *Ästhetik. Aufgabe(n) einer Wissenschaft*, Freiburg im Breisgau, S. 121–134
- Funke, Pia-Maria (1996): *Über das Höhere in der Literatur. Ein Versuch zur Ästhetik von Botho Strauß*, Würzburg
- Hammermeister, Kai (2007): *Kleine Systematik der Kunstfeindschaft*, Darmstadt
- Jaynes, Julian (1988): *Der Ursprung des Bewusstseins durch den Zusammenbruch der bikameralen Psyche*, Hamburg
- Kaden, Christian (2002): Art. „Musik. II. Ritualität in der Krise: Platons Musikphilosophie“, in: *Ästhetische Grundbegriffe*. Hg. von Karlheinz Barck u. a., Bd. 4: Medien – Populär, Stuttgart–Weimar, S. 261–263
- Kant, Immanuel (*KrV*): *Kritik der reinen Vernunft*. Hg. von Raymund Schmitt, Hamburg 1956 u. ö.
- Kishon, Ephraim (1986): *Picasso war kein Scharlatan. Randbemerkungen zur modernen Kunst*, München–Wien
- Krohn, Wolfgang, Hg. (2006): *Ästhetik in der Wissenschaft. Interdisziplinärer Diskurs über das Gestalten und Darstellen von Wissen*, Hamburg
- Maturana, Humberto R./Varela, Francisco J. (1987): *Der Baum der Erkenntnis. Wie wir die Welt durch unsere Wahrnehmung erschaffen – die biologischen Wurzeln des Erkennens*, Hamburg
- Menninghaus, Winfried (2001): Art. „Ekel“, in: *Ästhetische Grundbegriffe*. Hg. von Karlheinz Barck u. a., Bd. 2: Dekadent – Grotesk, Stuttgart–Weimar, S. 142–177
- (2003): *Das Versprechen der Schönheit*, Frankfurt am Main

- Peitsch, Helmut (2001): Art. „Engagement/Tendenz/Parteilichkeit“, in: *Ästhetische Grundbegriffe*. Hg. von Karlheinz Barck u. a., Bd. 2: Dekadent – Grotesk, Stuttgart–Weimar, S. 178–223
- Reicher, Maria Elisabeth (2005): *Einführung in die philosophische Ästhetik*, Darmstadt
- Städtke, Klaus (2001): Art. „Form“, in: *Ästhetische Grundbegriffe*. Hg. von Karlheinz Barck u. a., Bd. 2: Dekadent – Grotesk, Stuttgart–Weimar, S. 462–494
- (2003): Art. „Sprache der Kunst/Kunst als Sprache“, in: *Ästhetische Grundbegriffe*. Hg. von Karlheinz Barck u. a., Bd. 5: Postmoderne – Synästhesie, Stuttgart–Weimar, S. 619–641
- Voigt, Stefanie (2005): *Das Geheimnis des Schönen. Über menschliche Kunst und künstliche Menschen oder: Wie Bewusstsein entsteht*, Münster
- Welsch, Wolfgang (o. J.): *Philosophie und Kunst – eine wechselhafte Beziehung* [<http://www2.uni-jena.de/welsch/>, Zugriff: 01.02.2021]

Umsetzungen im Umweltdiskurs

Mülltonne, Speerschleuder und Fahrradschlauch – Über künstliche und natürliche Intelligenz

Dietrich Dörner

Zusammenfassung

Wenn Intelligenz die Fähigkeit bedeuten soll, „zwischen den Zeilen zu lesen“ (*inter legere*), dann stellt sich die Frage danach, wie dies gehen kann, bei natürlicher wie künstlicher Intelligenz. In Frage kommen Deduktion, Induktion und Abduktion. Sinnvolles deduktives Schließen setzt eine Menge von Hintergrundwissen voraus, bzw. das Bedürfnis, sich das jeweils erforderliche Hintergrundwissen zu beschaffen. Z.B. indem man danach fragt. Das überfordert die heutige künstliche Intelligenz. Induktives Schließen führt zu Verallgemeinerungen, aber nicht zu Begründungen. Diese lassen sich durch Abduktion gewinnen. Wie psychologische Modellbildung zeigt, sind dafür ein leistungsfähiges Protokollgedächtnis und sprachgestützte Fragefähigkeit wichtig. Zusätzlich ist es notwendig, bildhafte Denkinhalte, Vorstellungen also zu entwerfen und – entsprechend der Problemstellung – zu verändern.

Abstract

If intelligence is to be understood as the ability “to read between the lines” (*inter legere*), then the question arises how this can come to be, in natural as well as in artificial intelligence. Options are deduction, induction, and abduction. Meaningful deductive reasoning presupposes a lot of background knowledge or at least the “feeling” that some background knowledge is missing. This puts too much strain at least on contemporary artificial intelligence. Inductive reasoning allows for general-

ization, but not for explanations. These can be gained by abduction. As psychological models can demonstrate, this in turn is based on a powerful protocol-memory and the language-based ability to ask questions. Additionally, it is necessary to generate imaginations and to alter them according to the problem.

1. Einleitung

Neulich las ich in einer Zeitung, ich glaube, es war die FAZ, dass in Zukunft alle Routinearbeiten von „künstlicher Intelligenz“ (KI) erledigt würden. Das wunderte mich etwas, da ich Intelligenz eigentlich immer in Verbindung gebracht hatte mit der Fähigkeit, *neue* Aufgaben, *neue* Probleme bewältigen zu können, nicht aber Routineaufgaben. Wenn künstliche Intelligenz lediglich „Routine“ kann, dann ist ‚künstliche Intelligenz‘ eigentlich nicht Intelligenz. In diesem Artikel möchte ich mich ein wenig mit den alltäglichen Routine-Tätigkeiten befassen und mit der Intelligenz, die dafür notwendig ist. Und ich frage mich, was nun tatsächlich die KI beitragen kann zur Bewältigung der alltäglichen Routine.

Intelligenz kommt von *inter legere*, meint also Fähigkeit, *zwischen* den Zeilen lesen zu können. Eigentlich kann man zwischen den Zeilen gar nichts lesen, da ja da gar nichts steht. Intelligenz besteht nun darin, dass man das, was „zwischen den Zeilen“ eigentlich stehen sollte oder könnte, sich irgendwie ausdenkt. Unter „Lesen“ verstehen wir hier einfach „Wahrnehmen“. Man sieht etwas, einen Stein, ein Haus, eine Mülltonne oder sonst irgendwas. Und was heißt das nun?

Auf der Abbildung 1 sehen Sie eine Alltagsszene. Da steht eine Mülltonne. Und zwar auf der Straße! Dahinter steht ein Auto. Es ist blau! Usw. Wir sehen bestimmte Sachen, und wir wissen auch, aus welchen Teilen diese Dinge bestehen. Dabei kennen wir auch die verborgenen Elemente, die wir gar nicht sehen, wir wissen wie diese Dinge von innen aussehen, was man mit ihnen machen kann, welche Rolle sie für die Lebensführung spielen, usw. wir kennen die Bedeutung dieser Dinge. Das ist das, was man wahrnimmt. Und was steht „zwischen den Zeilen?“

Eine Methode, um das herauszubekommen, besteht darin, dass man aus den Dingen und den Relationen zwischen den Dingen etwas ableitet, etwas logisch deduziert. Es reicht nicht, dass man die Dinge kategorisieren kann, sondern man muss auch noch überlegen, was diese spezifische Konstellation bedeutet. Die Dinge sind

bekannt, aber die Konstellation ist immer anders. Und was kann man aus dem Auto, der Mülltonne, dem anderen Auto, dem blühenden Busch, usw. ableiten?



Abb. 1: Die Mülltonne.

Alle hier abgedruckten Abbildungen stammen vom Verfasser.

1. Dass die Tonne auf der Straße steht und nicht auf dem Parkstreifen ist gefährlich. Denn die Tonne könnte die Ursache eines Unfalls werden, da an dieser Stelle die Tonne die Straße verengt; zwei Autos könnten einander nicht mehr ausweichen.
2. Es ist Frühling, wie man an der blühenden Cornell Kirsche sieht. Im Frühling gibt es häufig böige Winde! Also könnte ein starker Windstoß die Tonne, die ja Rollen hat, leicht gegen das blaue Auto werfen. Und ich als Besitzer der Tonne müsste für den Schaden aufkommen. Und der Besitzer des blauen Autos könnte unter Umständen auf meine Kosten die gesamte Karosserie des Autos erneuern lassen. Und das würde teuer! Also: Die Tonne muss so schnell wie möglich hier weg und auf ihren Platz!
3. Oder man könnte ableiten, dass man mit dem Inhalt der Tonne (Altpapier!) gut und leicht ein Feuerchen machen könnte. Und es könnte nun jemand vorbeikommen, dem es einfallen möchte, sich doch ein wenig zu wärmen an diesem recht kalten Tag. Was könnte aber aus einem Feuer

entstehen, welches in der Nähe der Brennstofftanks von zwei Autos entfacht wird?

Also ganz klar: Die Mülltonne muss weg!

2. Deduktion

In der Logik sind Computer stark. Logische Spiele, wie z. B. Schach oder Go, spielen sie inzwischen viel besser als Menschen. Wie würde also nun eine „logische Analyse“ der Tonnen-Szene aussehen, wenn wir sie durch ein KI-System durchführen ließen?

Man kann nur das ableiten, was aus bestimmten Relationen zwischen den Dingen folgt. Eine KI könnte durch eine 3-dimensionale Interpretation des Bildes durchaus z. B. die räumlichen Abstände zwischen den Objekten, die da zu sehen sind, herausfinden. Sie könnte dann ableiten, dass die Mülltonne von dem blauen Auto etwa 3 m entfernt ist, das blaue Auto von dem davorstehenden schwarzen etwa 1,5 m. Nun gut; daraus aber könnte man noch nicht viel schließen.

Zusätzlich sollte die KI noch eine ganze Menge von *Physik* verstehen, um Ableitungen machen zu können. Sie müsste also beispielsweise die Materialeigenschaften der Dinge kennen und wissen, dass der Aufprall der Mülltonne auf das blaue Auto eine Beule der und der Tiefe und der und der Ausdehnung erzeugen könnte. Und dass beim Aufprall wahrscheinlich der Lack abplatzt. Der Lack vom Auto. Und es ist erfahrungsgemäß schwierig, eine Schadstelle im genau richtigen Farbton nachzulackieren. Vielleicht muss man das ganze Auto nachlackieren! Und das würde teuer! Außerdem müsste die KI etwas von *Klimatologie* verstehen, um feststellen zu können, dass ein böiger Wind im Frühjahr sehr wahrscheinlich ist. Dann könnte sie nämlich die Gefahren beurteilen, die dem blauen Auto durch die Mülltonne drohen. Denn die Mülltonne steht „westlich“ von dem blauen Auto; die Windböen aber kommen in unseren Breiten hauptsächlich aus dem Westen! (Man sieht die westliche Position der Mülltonne an der Neigung des Baumes hinter dem blauen Auto!)

Dann müsste die KI auch etwas von *juristischen* Zusammenhängen wissen, sie müsste nämlich wissen, dass der Eigentümer der Mülltonne verantwortlich für das Verhalten der Mülltonne ist, da die Mülltonne keine juristische Person ist. Und dass

er haftbar zu machen wäre für die Schäden, die die Mülltonne ihrer Umgebung zufügen kann, da ja die Mülltonne selbst nicht haftbar zu machen ist.

Weiterhin müsste aber die KI verstehen, dass der Eigentümer der Mülltonne höchstwahrscheinlich eine Abneigung dagegen hat, sein Geld in die Schandtaten seiner (nicht strafmündigen) Mülltonne zu investieren.

Schließlich müsste die KI noch wissen, dass eine *grüne* Mülltonne ein Depot für Papier- und Pappmüll ist, im Gegensatz zu einer *schwarzen* Mülltonne, in der Restmüll entsorgt wird, der nicht so gut brennt.

Und sie müsste noch etwas von *Psychologie* verstehen, nämlich vom Bedürfnis von Menschen sich – besonders in der kalten Jahreszeit! – zu wärmen, und die deshalb auf die Idee kommen könnten, zu diesem Zwecke mit den Pappstücken und Papierresten der grünen Mülltonne ein wärmendes Feuer zu entzünden. (Aber allerdings sollte eigentlich der KI auch noch einfallen, dass die Mülltonne doch wohl gerade entleert worden ist, sich vermutlich also derzeit keine Papp- und Papierüberreste in der Tonne finden lassen.)

Es steht zu fürchten, dass ein KI System mit dem Problem, zu entscheiden, was nun mit der Mülltonne zu geschehen hat, ziemlich überfordert wäre! Denn es ist mir kein KI System bekannt, welches zugleich über Kenntnisse in der Physik, der Klimatologie, der Juristerei, der Materialkunde verfügt, die man bräuchte, um ableiten zu können, welche Gefahren von der Mülltonne und ihrer Umgebung ausgehen könnten! Man könnte natürlich sagen: Das ist alles gut und schön, aber die KI soll gefälligst die Mülltonne wegräumen! Nun gut, so könnte man natürlich aus der KI einen braven Soldaten machen. „Tu, was Dir befohlen wird!“ Dafür aber braucht die „KI“ keine Intelligenz! Eine alte Unteroffiziersermahnung an die Soldaten lautet: „Sie sollen nicht denken! Überlassen Sie das den Pferden, die haben größere Köpfe!“

Auch könnte es sein, dass die Mülltonne mit Absicht auf die Straße geschoben wurde. Ein Bewohner des Hauses kam von einem Einkauf zurück, fühlte plötzlich ein Stechen in der Brust, fürchtete ohnmächtig zu werden, und gab deshalb der Mülltonne mit letzter Kraft noch einen Stoß, damit sie auf die Straße rollte und jedem passierenden Nachbarn klarmachen sollte: „Hier ist etwas in Unordnung! Du solltest einmal nachsehen!“ Nun würde einem Nachbarn das wohl nicht unbedingt einfallen, wenn er nur eine Mülltonne (die dahin nicht gehört) auf der Straße sehen würde. Wäre das aber verbunden mit anderen Indizien, z. B. mit offenstehenden Autotüren oder gar einem laufenden Motor, dann sollte der Nachbar ableiten können: „Hier stimmt etwas nicht!“ Und eine KI sollte ja wohl nun dasselbe leis-

ten können. Kann das die KI? Kurzum: Die KI mit ihren enormen Fähigkeiten im Go- und im Schachspiel wäre kaum in der Lage, mit dem Problem der Mülltonne, die auf der Fahrbahn steht, statt auf dem Parkstreifen, aufgrund logischer Schlüsse in adäquater Weise umzugehen! Weil sie das notwendige Hintergrundwissen nicht hat und weil wohl auch nicht das „Gefühl“ in der KI entsteht, dass ihr notwendiges Wissen fehlt.

3. Induktion

Neben dem logischen Schluss ist auch der induktive Schluss, die Verallgemeinerung, eine Form des Zwischen-den-Zeilen-Lesens. Da steht also ein Mülleimer auf der Straße, und zwar auf der Fahrbahn. Ist das nun ein Einzelfall? Oder stehen in der ganzen Stadt die Mülleimer auf der Fahrbahn. Dann wäre der Mülleimer vor meinem Haus nur ein Beispiel für eine sehr oft auffindbare Gefahrensituation. Gucken wir doch mal nach! Eine kurze Fahrt durch die Stadt zeigt, dass mindestens 60 % aller Mülltonnen nach der Leerung auf der Straße, konkret auf der Fahrbahn, stehen. Wir haben es also mit einem allgemeinen Missstand zu tun. Dem sollte man abhelfen. Dafür aber sollte man zunächst einmal ergründen, warum ist denn das der Fall? Und damit sind wir schon bei der nächsten Form des Lesens zwischen-den-Zeilen, nämlich der *Abduktion*. Der Begriff ‚Abduktion‘ stammt von Charles Sanders Peirce¹.

4. Abduktion

Nicht nur, was *folgen* könnte, ist in der Mülltonnensituation interessant. Interessant ist vielmehr auch noch, wie es denn eigentlich zu der Situation mit der Mülltonne gekommen ist? Oder wie es dazu hätte gekommen sein können? Die Abduktion besteht darin, dass man herausfindet, welches die Gründe dafür gewesen sind, dass die Konstellation so beschaffen ist, wie wir sie vorfinden. Die Abduktion bedeutet die Lösung der „Kommissarsfrage“. Kurz und brutal: „Hier ist die Leiche! Wie kam es dazu?“

¹ Peirce (1992).

Die vorgefundene Situation soll sich nun also als **Konklusion** einer zu erfindenden Schlusskette ergeben. Also: *Erfinden* wir die Gründe für die auf der Fahrbahn stehenden Mülltonnen!

1. Die Müllmänner haben es immer eilig. Und daher haben sie die Tonne, ohne viel zu überlegen, einfach wieder auf die Straße gestellt, nachdem sie entleert wurde. Die Müllmänner und ihre Motive sind der Ausgangspunkt eines Prozesses, der schließlich zu der „Tonnsituation“ geführt hat.
2. Neulich las ich in der Zeitung, dass es der Firma Auto-Meyer schlecht geht. Niedriger Umsatz! Könnte unter Umständen Auto-Meyer auf die Idee gekommen sein, die Müllmänner durch kleine Geldgaben dazu zu überreden, die Mülltonnen auf der Straße stehen zu lassen? Damit sich Unfälle ereignen und die Firma Auto-Meyer Reparaturaufträge bekommt? Und dadurch ihre finanzielle Situation verbessert!?

Könnte die KI in der gegebenen Situation solche Überlegungen durchführen? Natürlich müsste sie über elementare psychologische Kenntnisse verfügen; sie müsste nämlich über das Bestreben der Müllmänner möglichst schnell, besonders wenn die Witterungsverhältnisse nicht so angenehm sind, wieder ihre Dienststation zu erreichen, um sich mit einer Tasse heißen Kaffees aufzuwärmen.

Und außerdem müsste die KI jeden Morgen die Zeitung lesen, um über die finanzielle Situation von Auto-Meyer informiert zu sein. Und zugleich müsste die KI über die möglichen Pläne des Autohausbesitzers spekulieren können, mit Hilfe der Müllmänner den Auftragseingang zu erhöhen. Und wenn nun der Herr Auto-Meyer ein guter Freund des Besitzers der Mülltonne wäre, so müsste die KI, vertraut mit der sozialen Umgebung ihres „Administrators“ sich sagen: „Es kann ja wohl doch noch nicht im Interesse meines ‚Admins‘ sein, dass ich die Mülltonne hier wegräume und seinem Kumpel die dringend benötigten Einnahmen vermassele!“

Wie wahrscheinlich ist eine KI nun mit den sozialen Verhältnissen ihres „Besitzers“ vertraut und zugleich mit den Voraussetzungen und den Folgen, die eine Freundschaft mit sich bringt? Also: Die KI muss mir erst noch begegnen, die in der Lage ist, die verschiedenen Gründe dafür, dass Mülltonnen auf der Straße herumstehen, herauszufinden. – Übrigens: Eigentlich gehört auch die im letzten Abschnitt schon geschilderte Möglichkeit plötzlicher Herzbeschwerden zum Thema „Abduktion“.



Abb. 2: Der Fahrradschlauch.

Die Abduktion ist besonders dann wichtig, wenn einem etwas Unerwartetes begegnet. Wenn man z. B. plötzlich in eine Situation gerät, bei der die gewöhnlichen Verhaltensweisen, mit denen man die Situation bewältigen kann, versagen. Dann sollte man sich zu helfen wissen. Wenn etwas nicht so geht, wie es gewöhnlich gehen sollte, sollte man einen Ausweg finden können.

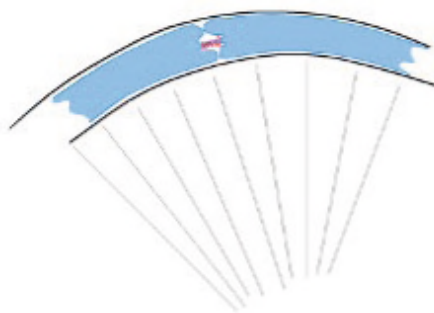


Abb. 3: Der Schlauch im Rad.

Nehmen wir einmal an, Sie befinden sich auf einer Fahrradtour und fahren auf einem Waldweg über eine knorrige Wurzel. „*Pschschtt*“ macht es und das Vorderrad ist platt! Ärgerlich! Aber Sie haben ja ihr Werkzeug dabei und deshalb ist das kein Problem. Eine kurze Inspektion der Satteltasche zeigt, irgendjemand hat das Flick-

zeug aus der Tasche entfernt. Nun stehen Sie also mitten im Wald mit einem „platten“ Fahrrad 8 km von ihrem Heim entfernt. Das Fahrrad auf die Schulter und dann ab nach Hause? Mühselig! Das Fahrrad schieben? Auch mühselig mit dem platten Reifen. Was tun? Abbildung 2 zeigt, was Sie tun können! Bei einer Wanderung ist es immer günstig, wenn man ein bisschen Material bei sich hat, um Unglücksfällen begegnen zu können. Man sollte vielleicht ein Pflaster dabei haben und einen Verband und ein Stück Draht, um irgendetwas zusammen zu halten, oder ein Stück Bindfaden, um etwas auf dem Gepäcksattel befestigen zu können. (Bei den Pfadfindern nennt man so etwas das AB Päckchen, „Allzeit-Bereit-Päckchen.“) Nun gut, dass Fahrradflickzeug war nicht vorhanden. Aber 2 m einer dünnen, starken Nylonschnur finden sich im AB-Päckchen. Als Flickzeug geeignet? Eigentlich ja nicht! Ja aber doch; man bindet doch Blutadern ab, damit das Blut nicht länger hinausfließt und eine Blutader ist doch auch eine Art von Schlauch. Warum soll man dann also keinen Fahrradschlauch „abbinden können“, so dass keine Luft mehr entweichen kann. – Gesagt, getan; das Ergebnis sehen Sie auf Abbildung 2. Über die Stelle mit dem Loch wurde, besonders vor und hinter dem Loch die Nylonschnur mit relativ großer Kraft gewickelt und ein Knoten krönte die Tätigkeit. Und Abbildung 3 zeigt, was dann wohl innerhalb des Schlauches geschah, als er wieder aufgepumpt wurde. Mehr als ein Jahr bin ich mit diesem Schlauch ohne Schwierigkeiten gefahren.

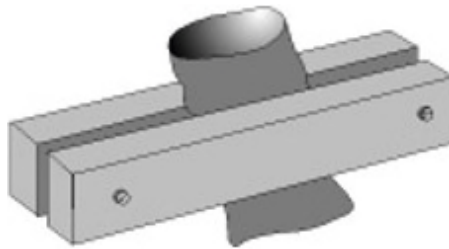


Abb. 4: Reparatur eines Fahrradschlauchs mit Hilfe von Klemmen. Natürlich braucht man zwei davon, eine vor dem Loch, eine andere dahinter. (Oder auch nur eine, wenn das Loch klein ist.) Und die Klemmen müssen viel kleiner sein!

Der Kern des Denkprozesses ist ein Analogieschluss, und das ist eine bekannte Denkfigur. „Ader = Schlauch!“ Könnte eine KI auch auf eine solche Idee kommen? Dazu müsste die KI über eine ganze Menge Erfahrungen verfügen, die nicht eigentlich zu dem Bereich „Fahrradreparatur“ gehören. Die KI müsste medizinische

Kenntnisse haben und wissen, dass man Adern, besonders Arterien, bei einer starken Verletzung abbinden muss und kann. Nun hat man aber sehr viele Dinge im Kopf, man weiß, dass man ein Brett auf ein anderes Brett aufschrauben kann. Und so könnte man beispielsweise auch den Fahrradschlauch „reparieren“, indem man vor und hinter dem Riss im Schlauch jeweils zwei Brettchen zusammenschraubt und so den Riss isoliert. Nun müssen aber die Brettchen so klein sein, dass sie in den Mantel des Rades hineinpassen und außerdem müsste man passende Schrauben haben usw.; das wäre also eine Methode des Flickens eines Lochs im Schlauch, die nur unter ganz bestimmten Umständen infrage kommt. Aber auch diese sollte einem einfallen können. Und dafür sind nun wieder andere Kenntnisse notwendig als für den Gebrauch der Nylonschnur. – Also: Eine KI so zu konstruieren, dass sie immer in der jeweiligen Situation die richtige Analogie findet, scheint mir nicht ganz einfach zu sein!

Bleiben wir noch ein bisschen bei den Problemen! Auf Abbildung 5 findet man drei sehr ähnliche Bilder. Es gibt jeweils drei Pflöcke, und auf dem ersten Pflöck liegen sieben runde kleine Holzscheiben. Das ist das Denkspiel „Turm von Hanoi“. Die Aufgabe besteht darin, den gesamten Turm von A nach C zu bringen; dabei darf nur jeweils eine einzige Scheibe bewegt werden und es darf niemals auf einen der Pflöcke eine größere Scheibe auf eine kleinere gelegt werden. (Überlegen Sie sich einmal die Lösungsmethode!)

Nun begegnet man Leuten, die, nachdem sie diese Instruktion gehört haben, sich keineswegs an die Arbeit machen, um ein Scheibchen nach dem anderen mal auf C, mal auf B zu legen, auch wieder auf A zurück. Nun gut, man darf niemals ein großes Scheibchen auf ein kleineres legen. Aber warum überhaupt die Scheiben verlegen? Man kann doch einfach die Pflöcke umbenennen. Man knibbelt das Etikett mit dem A einfach ab und dasselbe macht man mit dem C-Etikett. Und dann klebt man unter den ersten Pflöck das C-Etikett. Und schon steht der Turm auf dem Pflöck C, ohne dass man etwas bewegt hat! „Ja, emmh, ... aber so war das eigentlich nicht gedacht!“ sagt der Versuchsleiter. (Gewöhnlich fällt einer Versuchsperson, der diese „Lösungsabkürzung“ einfällt, auch sehr schnell die richtige Lösung ein.)

Eine andere Versuchsperson meint: „Also der rechte Turm soll nach links!?“ – „Genau!“ meint der Versuchsleiter. Die Versuchspersonen dreht das gesamte Arrangement einfach um 180°! Und meint zum Versuchsleiter: „So richtig?“ – Der Versuchsleiter: „Ja, emmh, ... aber so war das eigentlich nicht gedacht!“

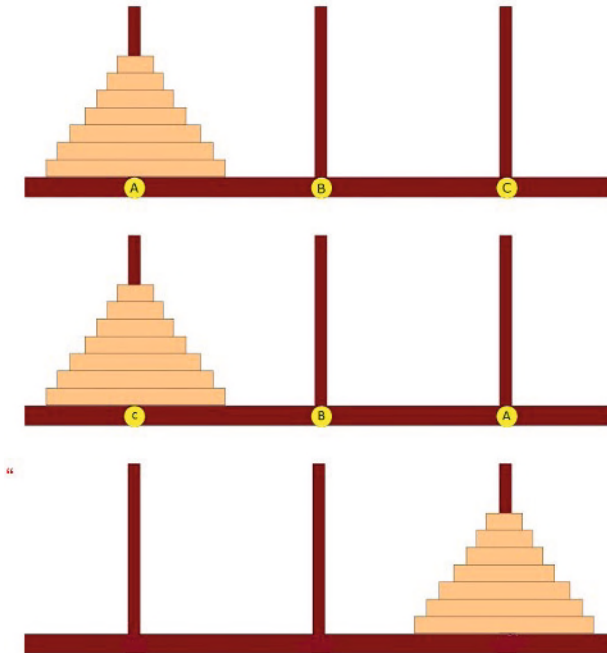


Abb. 5: Der „Turm von Hanoi“.

Natürlich wissen die Versuchspersonen ganz genau, dass sie eine andere Lösung suchen sollen. Aber sie sind einfallsreich und demonstrieren das und freuen sich dabei darüber, den Versuchsleiter in Verlegenheit gebracht zu haben. Demonstration der eigenen Überlegenheit! Sollte man sich nicht entgehen lassen! Darüber hinaus haben diese beiden „Lösungen“ ja durchaus eine philosophische Dimension: Was ist Bedeutung? Immerhin: Die beiden Versuchspersonen haben nicht gegen die Regeln verstoßen! Und das Ziel haben sie auch erreicht!

Kennt man irgendeine KI, der so etwas auch einfallen kann?

Kehren wir zum Ernst des Lebens zurück. Und jetzt wird es wirklich ernst. Auf der Abbildung 6 sieht man Neandertaler bei der Mammutjagd etwa im Jahre 80.000 v. Chr.² Sie brauchen Fleisch für die lange Winterzeit und so ein großes

² Anmerkung der Herausgeber: Abbildungen Nummer 6 („Neandertaler auf der Mammutjagd“) und 7 („Die Speerschleuder“) wurden zwar im Vortrag gezeigt, können hier

Mammut ernährt den ganzen Stamm für einige Wochen. Aber so ein Mammutbulle hat ekelhaft lange Stoßzähne und weiß sie zu gebrauchen und außerdem hatte er einen Rüssel, mit dem er einen der Mammutjäger durchaus erfassen kann, um ihn dann auf den Boden zu schmettern. Die Mammutjäger haben ihre Stoßlanzen; gefährliche Waffen mit rasiermesserscharfen Feuersteinspitzen auf einem starken Stab. Aber es hilft alles nichts: Sie müssen nah heran an das Mammut. Sehr, sehr gefährlich! Aber man braucht das Fleisch!

Die Neandertaler lebten (in West- und Mitteleuropa) über zehntausende von Jahren mit der moderneren Homo sapiens Bevölkerung, die unsere Ahnen darstellt, zusammen, den Sapiens-Menschen. Etwa gegen 20.000 v. Chr. starben die Neandertaler auch in Westeuropa aus; gewöhnlich wird das darauf zurückgeführt, dass sie sich an die „neuen“ Gegebenheiten nicht anzupassen wussten.

In diesem Zusammenhang ist es interessant, dass die modernen Homines sapients über eine Waffe verfügten, die den Lanzen der Neandertaler weit überlegen waren. Sie verwendeten „Speerschleudern“ (siehe Abbildung 7), gewissermaßen ein Hebelsystem, um Speere mit größerer Wucht und auf größere Entfernungen zu werfen. Mit der Speerschleuder war es möglich, Speere auf eine Entfernung von bis zu 150 m zu werfen. (Der augenblickliche Rekord im Speerwerfen (ohne Schleuder!) liegt bei etwa 100 m.) Durch diese Fernwaffe brauchten die sapiens-Menschen sich nicht mehr dem Mammut zu nähern und konnten außerdem eine Art von „Dauerfeuer“ auf das Mammut aufrechterhalten, da sie von diesen leichten Speeren 10 oder 12 mit sich führen konnten. Die Speerschleuder war also eine Art von steinzeitlichem „Maschinengewehr“; man konnte relativ schnell hintereinander viele Speere werfen.

Nun lebten die Sapiens-Menschen Zehntausende von Jahren mit den Neandertalern im selben geographischen Bereich, und es ist sicher, dass beide Menschenarten Beziehung zueinander hatten; unser Genom enthält etwa 2 % Neandertaler-Gene. Aber: Mögen Neandertaler und Sapiens-Menschen auch Tisch und Bett miteinander geteilt haben; die Neandertaler übernahmen die Speerschleuder als Waffe *nicht*. Warum nicht? Der Sapiens erfand noch viele andere Dinge, z. B. – ganz wichtig! – die Nähnadel. Dadurch konnten sich die Sapiens-Menschen Unterwäsche nähen

aus technischen Gründen aber leider nicht wiedergegeben werden. Wir bitten diese Auslassung zu entschuldigen, und meinen zugleich, sie verantworten zu können, da das Abgebildete im Text jeweils sehr anschaulich beschrieben wird.

und auch z. B. Anoraks und Fellschuhe. Dagegen umhüllten sich die Neandertaler mit Fellen, die sie vielleicht mit Riemen und Bändern an ihrem Körper befestigten. (Siehe Abbildung 6.) Die Sapiens Menschen erfanden auch das Kochen und die Töpferei, schnitzten Flöten, schufen Kunstwerke und entwickelten Religionen. Wahrscheinlich starben die Neandertaler aus, weil die Sapiens-Menschen einfach die besseren Jäger waren, mit den besseren Waffen, mit besserer Kleidung und Schuhwerk ausgestattet. Auch waren wohl die Sapiens-Menschen die besseren Organisatoren. Wie kam es, dass die beiden nahe verwandten Menschenarten sich so unterschiedlich entwickelten. Das Gehirn der Neandertaler war *größer* als das der Sapiens-Menschen (ca. 1.700 g; wir haben heute Gehirne mit einem Gewicht von 1.400 bis 1.500 g.)

Was war das Geheimnis der Sapiens-Menschen? Man nimmt heute an, dass vor etwa 70.000 Jahren eine „kognitive Revolution“ bei den Sapiens-Menschen stattgefunden hat, aber nicht bei den Neandertalern. In irgendeiner Weise muss sich das Gehirn der Sapiens-Menschen verändert haben. Harari (2017), der die Unterschiede zwischen Neandertalern und Sapiens-Menschen beschreibt, meint, dass wohl irgendeine Art von neuronaler Umprogrammierung in den Gehirnen der Sapiens-Menschen stattgefunden haben muss, weiß aber nicht, um welche Art von Neuverschaltung es sich da gehandelt hat.

Wir haben künstliche Lebewesen entworfen, die „Mäuse“. Dabei entdeckten wir, dass diesen Mäusen gleichfalls eine kognitive Revolution beschert wird, die sich von selbst einstellte, ohne dass wir danach gesucht haben. Unsere Mäusepopulation lebt in einer Gemeinschaft von etwa 100–125 Individuen; dann auch mal wieder 95 dann wieder 105 usw. auf einer Insel. Und es tut sich nicht viel. Die Mäuse haben Hunger, Durst, suchen Gemeinschaft, haben Sex, haben Freunde, Feinde. Sexuelle Fortpflanzung bedeutet eine Genrekombination, so dass eine große Verschiedenheit von Mäusen entsteht.

Wenn man das Mäuseprogramm eine Zeit lang laufen lässt, so dass viele Generationen vergehen, dann kommt irgendwann mit großer Sicherheit eine Phase, in der die Anzahl der Mäuse ungeheuer zunimmt; von etwa 100 springt die Population auf die Größe von 300, dann 400, dann bis zu 700 Mäuse an. Wir können natürlich feststellen, woran das liegt. Es liegt an einem längeren „Protokollgedächtnis“; das Protokollgedächtnis, das Gedächtnis für die unmittelbare Vergangenheit (das hippocampale Gedächtnis vielleicht?) wurde länger, dadurch, dass die Verfallsrate geringer wurde. Es war also ein ganz einfacher neuronaler Parameter, der sich

änderte. Aufgrund aber dieses längeren Protokollgedächtnisses konnten sich die damit begabten Mäuse längere Strecken ihrer Vergangenheit merken, wussten also auf einem Jagdzug, wann sie wo waren und auf welche Art und Weise sie an einen bestimmten Ort gekommen sind und das in allen Einzelheiten. Ihr Gedächtnis enthielt auf diese Weise einfach längere „Pläne“; die „Mäuse“ wussten in viel höherem Maße als der Neandertaler, wann Sie wo gewesen waren, wem sie dort begegnet sind, was sie begünstigte oder was ungünstig für sie war.

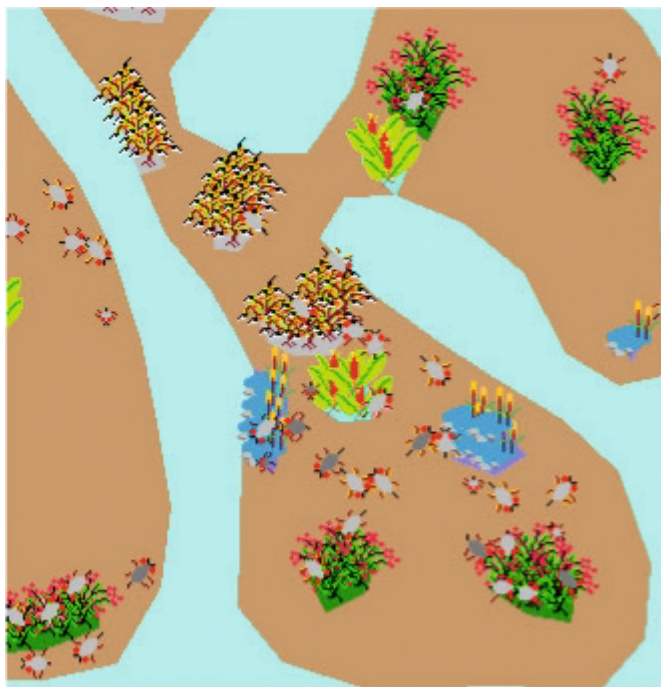


Abb. 8: Die Mäuse in ihrer Lebenswelt. Helle Mäuse: weiblich, dunkle Mäuse: männlich. Man sieht auch Kinder. Und außerdem Himbeerbüsche, Wasserstellen, Dornbüsche (gefährlich!), Heilpflanzen.

Und auf diese Art Weise hatten die Mäuse mit dem längeren Protokollgedächtnis einen viel besseren und viel weiterreichenden Überblick über das Gebiet, in dem sie lebten. Und sie erfuhren auch, dass es unter Umständen günstig war, nicht immer ganz genau den gleichen Weg zu einem Ziel zu gehen, sondern dass sie beispielswei-

se bei Regenwetter nicht einen bestimmten Abhang hinabgehen sollten, da dieser dann sehr rutschig war und man leicht ausgleiten und abstürzen konnte. Man sollte bei Regenwetter also diesen Berghang eher meiden. Ein anderer Weg war gefährlich, weil dort morgens immer ein Löwe auf Beute lauert; man sollte möglichst nur nachmittags dort vorbeigehen oder aber immer in größeren Gruppen, so dass man den Löwen vertreiben konnte. Es entwickelte sich also ein System von Alternativwegen, die bei bestimmten Bedingungen benutzt werden sollten oder zu vermeiden waren. Auf die Dauer entwickelte sich in einer „Maus“ ein relativ dichtes Netz von Wegen, die sich auch überkreuzten und Alternativwege enthielten. Verschiedene Teilstücke der Wege konnte man nun zu neuen Wegen zusammenfügen.

Jetzt hatte man also nicht nur ein größeres Wegenetz, sondern das Wegenetz war insgesamt viel flexibler mit Alternativwegen versehen. Wollte man nun jemand anderem den Weg schildern, so musste man Worte gebrauchen, die es vielleicht noch gar nicht gab: „Wenn ... dann ..., aber ...!“ „Wenn ... dann ... und nicht ..., dann dennoch ...!“ „Wenn aber ..., dann nicht ...!“ „Und obwohl ..., dann ...!“ „Wenn ..., dann ..., auch wenn nicht ..., dann trotzdem ..., aber ...!“ „Bevor ..., dann aber ..., wenn ..., sonst ...!“ „Vorher aber ...!“ „Wenn noch nicht ..., dann ..., obwohl ...dennoch ...!“ „Deshalb, zu dem Zweck ..., wenn aber ..., dann nachdem ...!“ „Unter der Bedingung, dass ...!“ Und da nun das gesamte Wegesystem in differenzierter Weise benutzt werden musste, je nach der Präsenz oder Absenz bestimmter Bedingungen, musste sich die Sprache entsprechend weiterentwickeln; es kam zu einer sehr differenzierten Grammatik. Auch lernte man, dass man in dem Wegenetz zu ein und demselben Ziel verschiedene Wege gehen konnte; auf diese Art und Weise wurden die einzelnen Wege in Teilstrecken unterteilt, die in verschiedener Weise kombiniert werden konnten.

Das machte zunächst die Bewegung viel flexibler. Aber diese Flexibilität kann sich auf andere Arten des Tuns ausgedehnt haben und man lernte, dass man auch anderes Tun entsprechend organisieren konnte. So beinhaltet ja die Speerschleuder statt eine Wurfbewegung zwei verschiedene, man wirft den Speer mit der Hand in eine Richtung, führt aber dann den Wurfstock nach und beschleunigt den Speer noch weiter. Also eigentlich werden hier zwei Aktionen hintereinander geschaltet.

Und es entwickelte sich auf diese Art und Weise die Fähigkeit, Kausalitäten zu kombinieren. Wenn A der Fall ist und auch B, dann musst du C machen, wenn aber B fehlt, dann kannst du C nicht machen, sondern du musst D machen. Die „Mäuse“ entwickelten ganz einfach eine differenziertere Vorstellung von den kausalen Zu-

sammenhängen in ihrer Umgebung und lernten es, Kausalitätsketten zu konstruieren. Und das war etwas, was die Neandertaler mit ihrem kurzen Protokollgedächtnis nicht konnten.

Das ist in groben Zügen die Theorie, die wir in Dörner (2021) genauer darlegen werden. Aufgrund ihres längeren Protokollgedächtnis waren die Sapiens-Menschen in der Lage, komplizierte Kausalzusammenhänge nicht nur auszunutzen, sondern auch zu konstruieren, so also die Speerschleuder (und später dann auch Pfeil und Bogen). Und sie erlebten *Geschichten*, nicht nur einzelne Ereignisse, und konnten einander diese Geschichten auch erzählen, also Pläne weitergeben. Und dadurch entwickelte sich die Sprache und mit der Sprache auch eben der freie Umgang mit Kausalitäten.)

Ich möchte annehmen, dass der Grund für den technischen Fortschritt, den die Sapiens Menschen machten, darauf zurückzuführen ist, dass sie in differenzierter Weise über Wege sprechen konnten. Und, dass sie nicht nur sagen konnten, was geschah und was dann geschah und was dann geschah, sondern dass sie ausdrücken konnten, *warum* etwas geschah, was es bewirkte, dass eine bestimmte Handlung einmal zu diesem Ende, ein anderes Mal zu jenem Ende führte. Sie konnten zwischen Ursachen und Wirkungen unterscheiden und auch differenzierte Ursachen benennen, also bedingte Ursachen, Ursachen, die nur dann wirksam werden, wenn bestimmte Bedingungen erfüllt sind. Und nachdem sie gelernt hatten, Ereignisse als Geschichten zu berichten, in denen dieses jenes erzeugt und jenes wieder etwas anderes und das noch mal etwas anderes, aber unter anderen Umständen ein drittes usw., erlernten sie es, die Welt als Ursache-Wirkungs-Netz zu betrachten und konnten aus diesem Grunde dann auch komplizierte Ursache-Wirkung-Geflechte gestalten und die Werkzeuge für die Prozesse erfinden.

Es genügte also vielleicht die Entwicklung des Protokollgedächtnisses, um zu komplizierten Werkzeugen zu kommen, die dann eben auch eine Optimierung der Handlungen möglich machten. (Das ist der Rahmen; es ist mir sehr klar, dass man das noch genauer beschreiben muss; am besten, man überlegt sich, wie sich das alles entwickeln konnte und zeigt durch Simulation, dass eben dies möglich ist.) Wenn diese Hypothese aber richtig ist, die wir soeben gerade grob geschildert haben, so kann man auch sagen, dass eine KI ohne eine differenzierte Sprache nicht möglich ist; man muss einen Prozess differenziert beschreiben können, und wenn man ihn nicht beobachten kann, so muss man es lernen, sich vorzustellen, wie das, was man nicht beobachten kann, vielleicht aussieht.

Also: Ohne die Sprachfähigkeit, genauer gesagt die Fähigkeit, sehr differenziert über Dinge berichten zu können, Kausalitäten und mögliche Qualitäten unterscheiden zu können, wird es niemals eine KI geben. Man darf nicht nur die Mülltonne sehen, sondern man muss den Kontext sehen, in dem sie steht, oder in dem sie stehen könnte, man muss wissen, welche Kräfte, Naturkräfte oder auch menschliche Motive, wirksam sein müssen, um eine Situation in einer bestimmten Weise zu gestalten.

Literatur

Dörner, Dietrich (2021): *Das verlorene Paradies. Unvollendetes Manuskript*

Harari, Yuval Noah (³⁰2018): *Eine kurze Geschichte der Menschheit*, München

Peirce, Charles Sanders (1992): *Reasoning and the Logic of Things*, Harvard

Corona und die Herausforderung für den Umweltschutz

Marion Friedrich/Joachim Rathmann

Zusammenfassung

Seit Jahren schon ist das Umweltbewusstsein in vielen Ländern hoch und das Wissen um umweltschädliches Verhalten genauso. Trotzdem bleiben Umweltschutzbemühungen oft defizitär. Die COVID-19 Pandemie führt zu starken individuellen Verhaltensänderungen und staatlichen Regulierungsmaßnahmen, in denen einige die Chance auf eine sozialökologische Transformation sehen. Auf individueller Ebene werden Mechanismen dargestellt, die eine solche Entwicklung hemmen und fördern können.

Abstract

For years, environmental awareness has been high in many countries, as has been the knowledge of environmentally harmful behavior. Nevertheless, environmental protection efforts often remain deficient. The COVID-19 pandemic is leading to strong individual behavioral change and governmental regulatory measure; in which some see an opportunity for socioecological transformation. At the individual level, mechanisms that can inhibit and promote such development are presented.

1. Einleitung

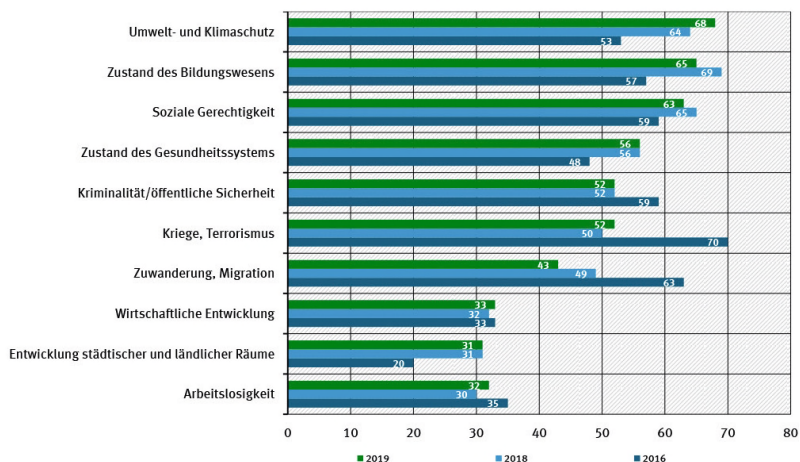
In den 2000/2010ern erlebte die Ökobewegung mit neuen Trends Aufwind, der sich bis 2019 fortsetzte: Natürlich ist dabei die Zunahme des ökologischen Bewusstseins kein Novum, sondern seit Jahrzehnten zu verzeichnen.

Aus einer philosophischen Perspektive heraus ist die erste der vier Kantischen Fragen: „Was können wir wissen?“ hinreichend beantwortet. Auch die zweite Frage: „Was sollen wir tun?“ ist geklärt. Und so rückt das Paradoxon immer mehr in den Vordergrund: Wenn wir seit geraumer Zeit wissen, welches Verhalten zu welchen Konsequenzen führt und wie wir mit einer Verhaltensänderung zur Minimierung der negativen Umweltfolgen beitragen können – weshalb tun wir es nicht?

Hier gibt es zahlreiche Erklärungsansätze, die vom Konzept der kognitiven Dissonanz nach Festinger über die Low-Cost-Annahme von Diekmann bis zum Einfluss des multioptionalen Konsumverhaltens, das Kuckartz beschrieb, reichen – und doch kann keine These einen allumfassenden Lösungsansatz erwirken.

Stellenwert des Umwelt- und Klimaschutzes

Rund zwei Drittel der Befragten schätzen im Jahr 2019 Umwelt- und Klimaschutz als eine sehr wichtige Herausforderung ein,



Angaben in Prozent
Repräsentativhebung bei etwa 2.000 Befragten pro Jahr
(2019: Zwischenerhebung, Online-Befragung)
Stichprobe ab 14 Jahren

Quelle: Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit/Unweltschutzesamt (Hrsg.),
Umweltbewusstsein in Deutschland 2019 - Ergebnisse einer repräsentativen Bevölkerungsumfrage

Abb. 1: Stellenwert des Umwelt- und Klimaschutzes. Quelle: <https://www.umweltbundesamt.de/bild/stellenwert-des-umwelt-klimaschutzes>, Zugriff: 05.07.2021.

In Deutschland wird Umweltschutz laut Umfragen mehrheitlich schon seit Mitte der 1980er als wichtiges Anliegen betrachtet und hielt bereits damals Platz 1 der Rangliste der aktuell bedeutsamsten politischen Probleme inne. 2016 hielten 53 % der Bevölkerung den Umweltschutz für eine wichtige politische Aufgabe, 2021 wur-

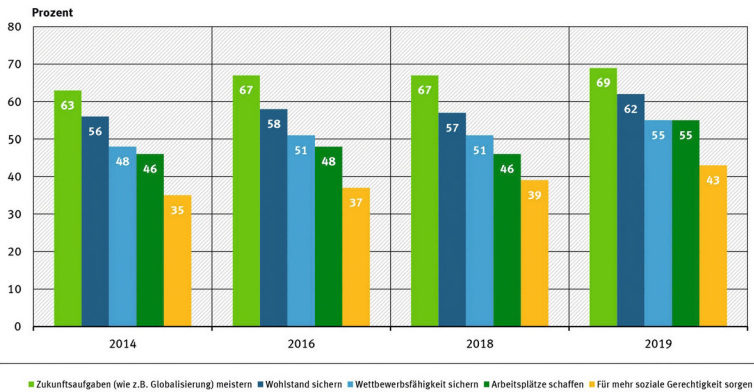
de Umwelt- und Klimaschutz noch vor Krieg und Terrorismus als sehr wichtige Herausforderung betrachtet (vgl. Abbildung 1, 2).

Und doch korrelieren Umweltbewusstsein und Umweltverhalten im Alltag nur gering miteinander. Umweltschutz wird zwar als gesellschaftsrelevantes Thema mit starken Auswirkungen auf die Bewältigung zukünftiger, z. B. Globalisierungs-, Aufgaben und Wohlstandssicherung betrachtet, doch ein hohes Umweltbewusstsein ist kein Indikator für konsequentes ökologisches Konsumverhalten.

Zeitvergleich: Umwelt- und Klimaschutz als Grundlage für politische Aufgaben

Frage: Welche Rolle spielt nach Ihrer Meinung der Umwelt- und Klimaschutz im Hinblick auf diese Aufgaben?

Antwort: Ein hinreichender Umwelt- und Klimaschutz stellt für diese Aufgabe eine grundlegende Bedingung dar.



Repräsentativhebung bei etwa 2.000 Befragten, 2. Befragungswelle, Stichprobe ab 14 Jahren, ohne „weiß nicht/keine Angabe“
Angaben für 2010 bis 2012: Basis: etwa 2.000 Befragte ab 18 Jahren
Angaben für 2014 bis 2019: Basis: etwa 2.000 Befragte ab 14 Jahren

Quelle: Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit/Umweltbundesamt (Bfug), Umweltbewusstsein in Deutschland 2019 - Ergebnisse einer repräsentativen Bevölkerungsumfrage

Abb. 2: Zeitvergleich Umwelt- und Klimaschutz. Quelle: <https://www.umweltbundesamt.de/bild/stellenwert-des-umwelt-klimaschutzes>, Zugriff: 05.07.2021.

2. Was bedingt die geringe Korrelation von Umweltbewusstsein und Umweltverhalten?

Nach dem „Erdgipfel“ der Vereinten Nationen (1992 in Rio de Janeiro abgehalten¹), bei dem auf die Wichtigkeit der Veränderung des Konsumverhaltens, um eine po-

¹ Vgl. https://www.bmz.de/de/themen/2030_agenda/historie/rio_plus20/umweltgipfel/index.html. (Zugriff 05.07.2021)

sitive Wirkung auf das Umweltbewusstsein zu erzielen, schienen die Themen der „Agenda 21“ eher wieder in den Hintergrund zu treten. Klar war bereits damals, dass eine Lösung der globalen Umweltprobleme nicht allein durch den Versuch der staatlichen Steuerung erreicht werden kann: Auch der Endverbraucher muss sich der Nachhaltigkeit seines Umweltverhaltens bewusstwerden. Diese Nachhaltigkeit im Sinne eines auf Verhalten wirkenden Umweltbewusstseins kann nicht auf Vorgaben und Bevormundung basieren, sie muss auf Freiwilligkeit, durch Einsicht bedingt, gründen.

Bereits im Umweltgutachten des Jahres 1978 wies der Rat der Sachverständigen auf Vollzugsprobleme im Umweltschutz hin², also auf die Schwierigkeiten, das Bewusstsein für kritisches Umweltverhalten in nachhaltiges Verhalten zu übersetzen.

Umweltwissen, Umwelteinstellungen, Umweltverhalten und Verhaltensintentionen gelten heute als verhaltensdeterminierende relevante Faktoren und damit zu differenzierenden Begrifflichkeiten. So führt Kuckartz³ aus:

- Umweltwissen beschreibt den Kenntnis- und Informationsstand einer Person über Umwelt und Natur, über Trends und Entwicklungen in ökologischen Aufmerksamkeitsfeldern.
- Unter Umwelteinstellungen werden neben Einstellungen zu Fragen des Umweltschutzes im engeren Sinne auch Ängste, Empörung, Zorn und Betroffenheit sowie persönliche Grundorientierungen und auf die Umwelt bezogene Werthaltungen verstanden.
- Mit Umweltverhalten wird das individuelle Verhalten in relevanten Alltagssituationen bezeichnet.
- Davon zu unterscheiden sind Handlungsbereitschaft und Verhaltensintentionen, Bekundungen, sich in Zukunft so und nicht anders verhalten zu wollen.

3. Auswirkungen der Corona Krise auf Umweltverhalten

Noch 2019 hätte nach Ansicht der Autoren ein umfassendes Bildungskonzept dazu beigetragen, die Korrelation zwischen Umweltbewusstsein und Umweltverhalten zu stärken. Hierzu wurde bereits 2020 von den Autoren ein Konzept erarbeitet, das

² Sachverständigenrat für Umweltfragen (1978).

³ Kuckartz (2013).

durch alle Bildungseinrichtungen, vom Kindergarten bis hin zu Universitäten, im multimodalen und interdisziplinären Vorgehen den abstrakten Wert „Nachhaltigkeit“ in erlebbares, ressourcenorientiertes Verhalten, auch unter Vermittlung von Suffizienzstrategien, abbildet. Strategien, die direkt und indirekt auf den Anreizwert nachhaltigen Umweltverhaltens Einfluss nehmen, umfassen beispielhaft:

- Logotherapeutische Wertearbeit: Erfahren von Bedeutsamkeit und Sinnhaftigkeit im Naturerleben
- Umbewertung der Umwelt in Mitwelt, um die Reziprozität und Interdependenz Mensch – Natur zu vergegenwärtigen
- Commitment-Vermittlung durch achtsamkeitsbasierte Ansätze
- Modelllernen und Eigenerfahrung als positive Verstärker
- verdecktes/imaginiertes Konditionieren durch Story Telling, positiv antizipierend wie aversiv, auch direkt erlebbar am Beispiel der Pflanzen im Schulgarten, die auf Umweltbelastungen reagieren und/oder Simulationsspiele
- Minimierung bzw. Auflösung der erlernten Hilflosigkeit⁴ zugunsten erlebter Selbsteffizienz⁵: Durch einfache verhaltenstherapeutische Methodik kann die Veränderungsmotivation gesteigert bzw. auch Frustrationstoleranz aufgebaut werden.

Als mögliche Faktoren für Unterlassungen umweltgerechter Handlungen identifizierten wir hier u. a.:

- fehlender/variabler Wertecodex und dysfunktionales Modellverhalten
- fehlende und/oder unzureichende Information, teils auch durch Informationsüberflutung bedingt: Ist die Bio-Karotte des Supermarkts, die aus Israel eingeflogen wird, wirklich „ökologischer“ als die heimische, aber nicht bio-zertifizierte Karotte vom Discounter?
- Insbesondere Jugendlichen ist es häufig aufgrund der sich noch in Entwicklung befindlichen synaptischen Verknüpfungen noch nicht möglich, valide Kostenfolgeabschätzungen und damit Zukunftsantizipation anzustrengen. Diese Fähigkeiten lokalisieren sich im präfrontalen Cortex, der ontogenetisch erst mit Ende der Pubertät ausreift

⁴ Vgl. das bereits in den 1970ern von Martin Seligman beschriebene depressogene Konzept der *learned helplessness*: Seligman (1975).

⁵ Vgl. Bandura (1977), S. 191–215.

- die Vermittlung hedonistischer und gegenwartsbezogener Prinzipien, im Sinne „man lebt nur einmal“
- Dystopische Zukunftsszenarien, im Sinne: „es ist eh zu spät“?

Das Bildungsprogramm zielt/e dabei darauf ab, das Verhalten IM ALLTAG umweltverträglicher zu gestalten. Nur, wenn ressourcen- und umweltorientiertes Verhalten so selbstverständlich geworden ist, dass es ritualisiert/automatisiert ausgeführt ist und keine kognitiven Entscheidungsprozesse mehr benötigt, wird nachhaltiges Verhalten als unanstrengend und „normal“ erlebt – was die unbedingte Voraussetzung ist, dass auch in kritischen und/oder Ausnahme-Situationen pro Mitwelt agiert werden kann. Natürlich erfordert dies wiederholtes Einüben der davor bewusst gewählten Verhaltensweisen: Die Einsicht muss ins praktische Handeln übersetzt und trainiert werden.

Und dann kam die Corona Krise: und damit eine deutliche Veränderung umwelt- und ressourcenorientierten Alltagsverhaltens. Als positive Auswirkungen der eingeschränkten Handlungsautonomie während des Lockdowns sind zu nennen:

- Umweltbelastungen sinken durch weniger Flug- und Fernreisen.
- Durch vermehrtes Homeoffice nehmen Straßenverkehr und damit einhergehende Umweltbelastungen ab.
- Die Wohnortumgebung wird durch Spaziergänge, Sport im Freien etc. zunächst entdeckt und zugänglich gemacht.

Als negative Auswirkungen des Lockdowns sind zu nennen:

- Produkte „to go“ werden erneut in Einmalverpackungen verkauft, wiederverwendbare Becher und Geschirr durfte von den Gastronomiebetrieben nicht mehr angenommen.
- Lieferservice-Bestellungen nahmen zu, damit steigt die Mülllast weiterhin.
- Supermärkte erhöhen ihren Lieferanteil.
- Internethändler erleben ein Hoch, damit einher geht die Zunahme von Umweltbelastungen durch Verpackung und Transport.
- Kochbox-Lieferungen mit ihren einzeln verpackten Single-Portionen sind gefragt wie nie.
- Homeoffice wirkt sich auf den Strom- und Energieverbrauch aus: Hier wird hochzurechnen sein, ob die gesparte Treibstoffbelastung dazu im Verhältnis steht.

- OP-Masken und FFP 2-Masken sind Wegwerfprodukte – die nicht immer im Restmüll landen, sondern häufig auf den Straßen und Spazierwegen „entsorgt“ oder verloren werden.

Es wirkt, als ob mit der empfundenen Bedrohung der eigenen Gesundheit und Existenz durch das Virus die Umweltachtsamkeit bei vielen Bürgern erloschen ist: Der Trend hin zu mitweltachtamen Verhalten oder gar Einsatz für unsere Mitwelt ist stark eingebrochen: Nicht nur, dass Versammlungen und Demonstrationen untersagt waren, Fridays for future haben an Aufmerksamkeit verloren.

Dies zeigt die bisherige kognitive Steuerung umweltschonenden Verhaltens deutlich auf. Es ist in der Psychologie keine neue Erkenntnis, dass erlebte Angst sich auf rationales und innovatives Verhalten auswirkt: Wer Angst hat, wendet vielmehr bereits erprobte und eintrainierte Verhaltensregeln an, es ist kein Raum für innovative oder rationale gesteuerte Handlungen.

Dies zeigt, dass erheblicher Nachholbedarf an der Übertragung abstrakter Werte wie Umweltbewusstsein in konkretes Alltagsverhalten wie ressourcen- und mitweltorientiertes Leben vonnöten ist: Natur muss als Mitwelt erfahrbar sein, damit wir Menschen auch in ambivalenten Situationen in der Lage sind, unser Überleben nicht unabhängig von eben jener Mitwelt in den Vordergrund zu stellen.

4. Ansätze zur Erlebbarkeit der Umwelt als Mitwelt

Durch die unfreiwillige Isolation infolge der politischen Corona-Auflagen verliert der Mensch mehr und mehr die Stabilität der „Herde“: Gruppenzugehörigkeit und damit einhergehendes soziales Feedback wirkten stabilisierend auf die Psyche und damit auf die erlebte Zufriedenheit im Leben. Eine „Herde“ in einer virtuellen Konferenz kann dies nicht ersetzen.

In der durch Martin Seligman mit begründeten positiven Psychologie werden die bereits in der Logotherapie von Viktor Frankl beschriebenen drei Wege zu einem erfüllten Leben paraphrasiert. Viktor Frankl nennt drei Lebensformen, mit denen der existentiellen Frustration des vergebens sinnsuchenden Individuums entgegnet werden kann⁶:

⁶ Frankl (1978), S. 81.

- Der *homo faber* verleiht seinem Dasein durch seine Arbeit Sinn.
- Der *homo amans* empfindet Bedeutungsvolles in Liebe und Sorge.
- Der *homo patiens* vermag es aufgrund der Trotzmacht des Geistes auch Leiden Sinn zu geben.

Während die Determinanten im Leben des *homo faber* Erfolg und Misserfolg sind, sind Erfüllung und Verzweiflung die bestimmenden Faktoren des *homo patiens*, wobei die Wertrichtungen des *homo faber* und *patiens* voneinander unabhängig sind, Erfüllung ist nicht abhängig vom Erfolg. Insbesondere in Krisensituationen kommen *homo faber* und *homo amans* ans Limit, während *homo patiens* die immer bleibende Freiheit der Stellungnahme zu seinem Schicksal, mithin die Möglichkeit der Wahl, wahrnimmt.

Übertragen auf unsere Thematik meint das, dass auch in Zeiten einer Pandemie diese Möglichkeit der Stellungnahme zum erlebten Unkontrollierbaren nicht nur bleibt, sondern fast verpflichtend für alle Menschen gilt: Dies ist der Weg raus aus der erlernten Hilflosigkeit, raus aus der passiven Opferrolle des sich fügenden Leidenden. Erst 2016 wurde das Konzept der erlernten Hilflosigkeit von M. Seligman selbst und S. Maier insofern adaptiert, dass nun Passivität als Reaktion auf einen Schock als standardmäßige, ungelernte Reaktion auf längere aversive Ereignisse beschrieben ist⁷.

Wenn also das Lebewesen über einen längeren Zeitraum hinweg nicht kontrollierbaren Situationen ausgeliefert ist, reagiert es mit Passivität: Dies ist neurobiologisch gut erforscht. Dieser Prozess, der auf das Überleben des Organismus abzu zielen scheint, ist serotoningesteuert. Der dorsale Nucleus Raphe unterdrückt so Flucht- und Wehrimpulse. Die Raphe-Kerne sind Teil der *Formatio Reticularis* und befinden sich im Hirnstamm – der Hirnbereich, der zahlreiche Vitalfunktionen reguliert.

Durch das Erlernen von Selbstwirksamkeit (und Psychotherapie zielt ja immer auch darauf ab, genau diese Selbsteffizienz wiederherzustellen: Also Wahlmöglichkeiten zu vermitteln, sich so oder eben anders zu verhalten!) kann diese Passivität überwunden werden. Das Ausüben und Erleben von Kontrolle im eigenen Leben gehen mit der Aktivierung des medialen präfrontalen Cortex einher, der seinerseits dann den dorsalen Raphe-Kern inhibiert.

⁷ Vgl. Maier/Seligman (2016).

Wenn Hilflosigkeit erlebt wird, treten Interessenverlust, Antriebsminderung, Motivationslosigkeit auf. Austausch, Kommunikation wird dann nicht mehr als entlastend, sondern als verpflichtend und belastend erlebt.

Obgleich die Förderung von Kohärenz und Selbstwirksamkeit seit vielen Jahren auch vom Gesundheitsministerium als wichtige Maßnahmen zur Krankheitsprävention und Gesundheitsvorsorge proklamiert wird, werden diese wichtigen psychohygienischen Faktoren aktuell vernachlässigt. Es gilt, vom sich Fügen und Ergeben wieder ins Handeln zu kommen. Insbesondere in der Verantwortungsübernahme für sich, sein Handeln, und für die Mitwelt kann es gelingen, vom Erdulden ins Bewirken zu kommen: So können sich *mind, body and action*, die drei Grundpfeiler eines integrierten Selbst, im *meaningful life*, im *bedeutungsvollen Leben*, verbinden.

Da nicht die ego-zentrierte Selbstverwirklichung zum Lebenssinn gemacht wurde, sondern im Umkehrschluss Selbstverwirklichung allein als die Folge eines sinn-erfüllten Lebens betrachtet wird, sind wir unabhängig(er) von Externa und Señor Rossis Suche nach dem letzten Faktor, der endlich das Glück bewirken soll. Auch in extremen Notlagen kann uns eine praktische Sinngebung beim Überleben helfen. Sinn im Leben können wir dabei durch

- die Hingabe an eine Sache (der schöpferische Wert)
 - die Hingabe an einen anderen Menschen (der emotionale Wert)
 - das würdige Ertragen eines Leids (der existenzielle Wert)
- erfahren.

In Seligmans Theorie ruht Wohlbefinden auf fünf Pfeilern, deren Anfangsbuchstaben sich in der Abkürzung PERMA wiederfinden:

- Positive Emotions („positive Emotionen“)
- Engagement („Engagement“)
- Relationships („Beziehungen“)
- Meaning („Bedeutung“/„Sinnhaftigkeit“)
- Accomplishment („Leistung“, umfasst das Konzept der erlebten Selbstwirksamkeit)

Nicht allein das Erleben möglichst zahlreicher Glücksmomente im Leben und die Möglichkeit bzw. die Fähigkeiten, diese selbst zu bestimmen, was das *pleasant life*,

das freudige Leben, widerspiegeln würde, bedingt stabiles menschliches Wohlbefinden: Das so genannte *good life*, das gute Leben, basiert bereits auf gesellschaftlichen und individuellen Werten und meint, Arbeit, Freundschaften, Liebesbeziehungen, Freizeit etc. selbst zu gestalten. Das *meaningful life*, das bedeutungsvolle Leben, reicht weit darüber hinaus: Dadurch, dass die eigenen Fähigkeiten und Kraft in den Dienst von etwas, was mein eigenes Sein transzendiert, erfahre ich Sinn und Bedeutung in meinem Leben.

Hier handle ich absichtsvoll und bedeutungsbewusst. Übertragen auf Franks Konzept der existentiellen Frustration und unser Konzept der Mitwelt kann ich die Bedeutung von Leben, auch meines Lebens, durch die direkte gefühlte Einbindung meines Ichs in die mich umgebende und bedingende Natur, bestehend aus meinen Mitmenschen wie meiner Mitwelt, erspüren: Jenseits von wohl durchdachten und damit von einem hohen Energieaufwand begleiteten Handlungsweisen, die sich ausgesprochen ambivalenzinstabil zeigen, kann durch konkretes TUN der angstbedingten Egozentrik auch in globalen Krisenzeiten entgegengewirkt werden.

Dieses Tun, wenn es alltagstauglich sein soll und in wohltuende Routinen führen soll, darf dabei nicht zu ambitioniert sein: Auch, wenn es darum geht, aus der erlernten Hilflosigkeit heraus zurück in Selbstwirksamkeit und Handlungsautonomie zu finden, ist es wichtig, kleinschrittig vorzugehen.

Insbesondere dann, wenn soziale Anbindung als eine der wichtigsten Ressourcen fehlt, ist Naturerleben hinsichtlich Resilienz und Frustrationstoleranz umso bedeutender. Wir Menschen als soziale Tiere sind ausgesprochen traumaresistent, wenn wir uns als Gruppenmitglied in einem sicheren sozialen Umfeld befinden: Bevor überhaupt die klassischen animalischen Überlebensreaktionen des Fight-Flight-Freeze in Kraft treten, suchen wir Unterstützung und Orientierung in unserer peer group. Dieser Mechanismus, der höheren Säugetieren angeboren ist, beschreibt Stephen Porges in seiner Polyvagal-Theorie⁸. Wenn jedoch diese Sicherheit fehlt, weil entweder von unseren Bezugswesen selbst Gefahr ausgeht oder sie nicht verfügbar sind, wenn Angriff und Flucht vor der Gefahr nicht möglich sind (ähnlich die bei der medial aufbereiteten Bedrohung durch ein Virus), greift als letzter Überlebensversuch die Erstarrung. Doch das subjektive Gefühl der Anbindung kann auch im Naturerleben erzielt werden: Sich als Teil des Ganzen, Teil von etwas Übergeordneten wahrnehmen, kann auch z. B. im Wald gelingen.

⁸ Porges (2010).

Wenn Trauma also ein Einfrieren von (unaushaltbaren) Gefühlen meint, meint seelisches Heilen ein Auftauen der Gefühle. So können sie nach und nach Teil meines Selbst werden, sie können in mein Ich reintegriert werden. Dabei ist die Wahrnehmung der eigenen Empfindungen und Emotionen Grundvoraussetzung für die Wahrnehmung und Einordnung der Emotionen anderer. So lange jemand also seine eigenen Gefühle nicht kennt oder unterdrücken muss, kann er nicht empathisch mit anderen sein. Darüber hinaus tendiert gerade der entwicklungs-traumatisierte Mensch dazu, mimische Expressionen anderer Menschen zu missinterpretieren: So wird ein neutrales Gesicht beispielsweise als bedrohlich oder auch eine angespannte Mimik als positiv erregt wahrgenommen.

5. Walderleben

Trockenstress in den vergangenen Jahren sowie Sturmereignisse haben eine massive Schädigung auf die Wälder nicht nur in Bayern entfaltet. Neben der Holzproduktion und weiteren ökosystemaren Leistungen dienen insbesondere urbane Wälder der Erholung und Gesundheitsförderung, so dass Wald als Gesundheitsressource inzwischen zu einem Thema geworden ist, das zunehmend in der Öffentlichkeit aber auch in der Politik thematisiert wird. So gibt es vermehrt Angebote zu verschiedenen Formen einer Waldtherapie, und seit 2016 gibt es einen ersten Therapie- und Kurwald auf Usedom⁹. Mit den Schlagworten Waldbaden und Waldtherapie haben Waldbesuche eine große Popularisierung erfahren. Das aus Japan stammende „Shinrin Yoku“, die japanische Bezeichnung für Waldbaden, bedeutet „ein Bad in der Atmosphäre des Waldes nehmen“. Dieses Naturerleben steht nicht nur mehr und mehr auch wegen zeitweise geschlossener Fitnessstudios im Trend, sondern ist auch Gegenstand zahlreicher Studien geworden: So publizierten z. B. Mathew P. White vom European Centre for Environment and Human Health, University of Exeter Medical School, Exeter, UK, und Kollegen im Jahre 2019 in *Scientific Reports*¹⁰ ihre Studie, die besagt:

⁹ Schuh/Immich (2019).

¹⁰ *Scientific Reports* 9, Article number: 7730 (2019).

A growing body of epidemiological evidence indicates that greater exposure to, or 'contact with,' natural environments (such as parks, woodlands and beaches) is associated with better health and well-being, at least among populations in high income, largely urbanised, societies. While the quantity and quality of evidence varies across outcomes, living in greener urban areas is associated with lower probabilities of cardiovascular disease, obesity, diabetes, asthma hospitalisation, mental distress, and ultimately mortality, among adults; and lower risks of obesity and myopia in children. Greater quantities of neighborhood nature are also associated with better self-reported health, and subjective well-being in adults, and improved birth outcomes, and cognitive development, in children.

Kurz: Die sympathische Nervenaktivität, die sich unter Stress erhöht, verringert sich beim Waldbaden mit weitreichenden Auswirkungen auf den ganzen Organismus, u. a.:

- Die parasympathische Nervenaktivität, die bei Entspannung steigt, nimmt zu.
- Die Stresshormone Adrenalin, Noradrenalin und Cortisol sinken.
- Der Blutdruck sinkt.
- Die Pulsfrequenz sinkt.
- Das Immunsystem wird gestärkt

Die positiven gesundheitlichen Effekte von Waldbesuchen gehen über eine kurzfristige des individuellen Wohlbefindens hinaus, weil durch die Blutdruckreduktion das Risiko für Herz- und Gefäßerkrankungen insgesamt gemindert werden kann. Waldbaden ist in der japanischen und südkoreanischen Tradition als ein multisensorisches Walderleben angelegt¹¹. In langsamen Spaziergängen, unterbrochen von Pausen werden unterschiedliche Sinne angesprochen¹²:

- In Wäldern ist die direkte Sonneneinstrahlung gedämpft und führt zu einem ständig wechselnden Spiel von Licht und Schatten. Visuelle Eindrücke variieren daher stark und unterstützen eine Form der ungerichteten Aufmerksamkeit, welche zur Entspannung beiträgt¹³.

¹¹ Hansen u. a. (2017).

¹² Vgl. Schuh/Immich (2019), S. 14f.; Rathmann (2020b).

¹³ Kaplan und Kaplan (1989).

- Die olfaktorische Wahrnehmung wird über natürliche Gerüche angeregt. Insbesondere bei Nadelhölzern nach Regenereignissen oder bei Trockenheit machen sich volatile Terpene positiv bemerkbar¹⁴.
- Wälder sind, verglichen mit einem urbanen Umfeld, ruhiger Räume. Naturnahe Geräusche treten in den Vordergrund und tragen zu einer Verbesserung der individuellen Stimmung bei.
- Das Essen von Waldbeeren oder Wildpflanzen spricht den Geschmackssinn an.
- Taktile Sensoren lassen sich über das Berühren von Baumrinden, Moosen oder Blättern stimulieren.

Waldbesuche können folglich eine therapeutische Wirksamkeit entfalten und der Stabilisierung von Gesundheit dienen sowie generell das Stresserleben reduzieren. Waldtherapie lässt sich somit als eine präventive Therapie verstehen, die die Wahrscheinlichkeit des Eintretens von Krankheiten zu verringern sucht. Darüber hinaus haben Wälder oft eine starke spirituelle Bedeutungszuschreibung. Spirituelle Werte beschreiben das Gefühl, mit etwas Anderem in Verbindung zu treten. Dieses Andere übersteigt das eigene Selbst und entwickelt eine große Bedeutung für die individuelle Sinndimension. Durch diese Form der Selbsttranszendenz und der eingehenden Verbindung mit anderen geistigen Wesen, der Natur oder Mitmenschen, erfährt das Selbst eine Vergrößerung gegenüber dem bloßen Ich. Diese gleichsam irrationale Gewissheit, am Anderen teilzuhaben, versucht Rudolf Otto zu rationalen Aspekten ins Verhältnis zu setzen, in dem er das „Irrational-Numinose“¹⁵ durch rationale Begriffe schematisiert. Damit versucht er den Begriff des „Heiligen“ zu erfassen. Das Gefühl der „Kreatürlichkeit“ beschreibt er dann mit dem Numinosen¹⁶. Als Moment des Numinosen stellt er das *mysterium tremendum et fascinosum*, das Gefühl des „schauervollen Geheimnisses“¹⁷ dar. Neben dem Schauervollen, das gleichsam den alttestamentarischen Gottesschrecken beschreibt (vgl. 1. Mose 35, 5), ist das Gefühl von Ekstase und Verzückung gestellt, das „Dionysische“¹⁸. Dieser Doppelcharakter des Numinosen lässt sich auch auf viele Bereiche der Naturerfah-

¹⁴ Cho u. a. (2017).

¹⁵ Otto (2014), S. 61.

¹⁶ Ebd., S. 8ff.

¹⁷ Ebd., S. 13.

¹⁸ Ebd., S. 43.

rung beziehen, denn dabei lassen sich auch Momente großer Faszination, Begeisterung und Abschreckung erfahren. Waldbesuche, verbunden mit entsprechenden Ritualen, lassen spirituelle Erfahrung in besondere Weise durchleben. Wälder weisen eine sehr lange spirituelle Bedeutungszuschreibung auf, die sich schon in der Antike an heiligen Hainen zeigte. In vielen Teilen Mittel- und Nordeuropas sind dunkle Haine, insbesondere mit Eichen und deren Mistelzweige, als heilige Stätten keltischer Druiden belegt. Später wurden durch christliche Klöster, Bäumen und Wäldern religiöse Zuschreibung beigelegt, denn die Wälder dienten nicht nur als Einnahmequelle für Klöster, sondern auch als Orte der Verehrung, um in der Abgeschiedenheit ein gottgefälliges Leben zu führen.

Auf einer globalen Skala gibt es weiterhin eine starke Verbindung von Spiritualität – Natur und menschlicher Gesundheit. So lässt sich monieren, dass der westliche biomedizinische Gesundheitsbegriff zu kurz greift¹⁹. Denn menschliches Wohlbefinden ist eingebettet in eine „gesunde“ Natur, da degradierte Ökosysteme letztlich auch die menschliche Gesundheit negativ beeinflussen können. Die Zerstörung unberührter Lebensräume kann neue Übertragungswege für Zoonosen eröffnen. Weitere Pandemien können daraus erwachsen. 2016 forderten Vertreter autochthoner Völker und medizinische Anthropologen in einem offenen Brief an die WHO, den westlichen Gesundheitsbegriff zu überdenken und zu erweitern, da ein Gleichgewicht der Menschheit mit seiner natürlichen Umgebung zentral sein muss, für ein umfassendes Verständnis von Gesundheit²⁰. Hier greift auch eine spirituelle Perspektive, denn Schäden, wie wir sie der Natur zufügen, schaden uns insofern selbst, als wir Teil dieser größeren Entität sind. Ein Leben im Gleichgewicht mit der Natur sowie das Ausleben der eigenen Spiritualität bilden daher zwei weitere Pfeiler eines umfassenden Gesundheitsbegriffs. Denn

aus animistischer Perspektive liegt der Ursprung all dieses physischen und psychischen Leids in der Gewalt, die unsere Zivilisation dem Ökosystem des Planeten ohne Not zufügt; nur wenn wir letztere beenden, werden wir erstes heilen können²¹.

¹⁹ Vgl. Rathmann (2020a).

²⁰ Charlier u. a. (2017).

²¹ Abram (2015), S. 44.

In der von Antonovsky vorgestellten salutogenetischen Perspektive von Gesundheit ist Sinn eine zentrale Widerstandsressource, um den Wechselfällen des Lebens begegnen zu können. Das Kohärenzgefühl beschreibt, wie ein Individuum belastende Umweltreize bewertet und durch Rückgriff auf verschiedene Ressourcen diesen begegnet. Die Bedeutung von Sinn bei der Bewältigung von Stressoren unterstreicht Antonovsky im Rückgriff auf Viktor Frankl. Sinn wird somit zu einer bedeutsamen Widerstandsressource bei der Aufrechterhaltung und Stärkung von Gesundheit. Gerade bei starken Belastungen erweist sich Spiritualität als eine zentrale Sinnressource²², deren Bindungskraft stärker ist als Sinngebung, die über die Familie, die Arbeit oder die Freizeit erfolgt. Sinn, der eine Ewigkeitsperspektive, die Zugehörigkeit zu Anderem eröffnet, kann gerade in schwierigen Lebensphasen eine große entspannende Wirkung entfalten. Hierin unterscheiden sich Formen der Spiritualität aus dem Wellness-Bereich, die gerade in westlichen Kulturen auf eine große Aufnahmebereitschaft stoßen, von älteren Religionen, weil das Glück ausschließlich im irdischen individuellen Hier und Jetzt gesucht wird. Zusätzlich lässt sich zeigen, dass ein religiöser oder spiritueller Lebensstil vielfach positive gesundheitliche Wirkungen entfaltet. Denn dies geht meist mit einem in vieler Hinsicht gemäßigteren Lebensstil einher, Gemeinschaftserfahrung und meditative Momente stärken weiter die physischen, psychischen und sozialen Aspekte von Gesundheit²³.

Affektive und emotionale Bedeutungszuschreibungen an die umgebende Natur weisen daher, jenseits eines rein naturwissenschaftlichen Naturzugriffs, auch über die Sinndimension eine wichtige Bedeutung für menschliches Wohlbefinden. Durch solche individuellen Raumaneynungen werden Orte qualitativ durch Bedeutungszuschreibungen konstruiert. „Sense of place“ beschreibt solche subjektiven Bindungen an Ort als Interaktion mit der physischen Umwelt können diese ausgleichend auf einen Menschen wirken und eine Stressreduktion unterstützen²⁴.

Die Verringerung des Stresserlebens wirkt sich wiederum direkt und indirekt auf die Empathiefähigkeit (als Mitgefühl) aus. Empathie wird hier nicht mit Mitleid, das durch das Überflutetwerden von Gefühlen einer anderen Person, gekennzeichnet ist, gleichgesetzt: Während Mitleid sich eher negativ auf die Selbst-

²² Bucher (2011), S. 127f.

²³ Vgl. ebd., S. 100ff.

²⁴ Korpela (1989).

wirksamkeit auswirkt, eher handlungsunfähig macht, trägt Empathie dazu bei, die Situation verbessern zu wollen, also wahrgenommenes Leiden zu verringern. Dabei sollte Empathietraining mit Wertearbeit synchronisiert werden, um dem gruppenpsychologischen Effekt vorzubeugen, dass bei zunehmenden empathisch motiviertem Verhalten häufig empathischer („wohlwollender“) mit Mitgliedern der eigenen Gruppe und zu Lasten von Außenstehenden umgegangen wird. Um empathisch sein zu können, ist die Wahrnehmung eigener Gefühle dabei Voraussetzung²⁵. Am Empathieerleben beteiligte Hirnregionen, z. B. die Insula, sind bei alexithymen²⁶ Personen nur mangelhaft aktiviert. Wer unfähig ist, seine eigenen Gefühle zu erkennen, hat damit auch Schwierigkeiten, Empathie gegenüber anderen Menschen zu entwickeln. Die Fähigkeit, innere Körperzustände angemessen zu identifizieren, ist hierbei ausschlaggebend – und dieses Körpererleben kann im achtsamkeitsbasierten Naturerleben unterstützt werden.

Natur- und Walderleben insbesondere wird in der *Scientific Community* mehr und mehr als auf prosoziales Verhalten positiv wirkender Faktor beschrieben. Gerne wird dabei auf das Wort von Immanuel Kant, mit dem er seine Kritik der Praktischen Vernunft abschloss, zurückgegriffen, um die Verbindung von Natur und wohlwollenden, empathischen und damit ethisch akzeptierbaren Einstellungen und Handlungen aufzuzeigen:

„Zwei Dinge erfüllen das Gemüt mit immer neuer und zunehmender Bewunderung und Ehrfurcht, je öfter und anhaltender sich das Nachdenken damit beschäftigt: der bestirnte Himmel über mir und das moralische Gesetz in mir.“²⁷ Ein zwei- bis dreistündiger, gemächlicher Waldspaziergang soll sich direkt empathieverstärkend auswirken²⁸ und die Probanden dazu veranlasst haben, in Experimenten teilungswilliger, rücksichtsvoller und weniger egoistisch vorzugehen. Diese Annahme wird dadurch unterstützt, dass Empathiefähigkeit mit der Fähigkeit, eigene Empfindungen und Gefühle wahrzunehmen, einhergeht²⁹: Naturbegegnungen wirken sich durchaus auf die eigene Körperwahrnehmung und damit auf das Gewahrsein von Empfindungen und Emotionen aus.

²⁵ Vgl. Bird (2010).

²⁶ Alexithymie: Gefühlsblindheit, meint das Unvermögen, eigene und fremde Gefühle wahrzunehmen und zu erkennen.

²⁷ Kant (1977), S. 300.

²⁸ Vgl. Spitzer (2009).

²⁹ Vgl. Bird u. a. (2010).

Doch darüber hinaus rechnen wir mit dem erwünschten Nebeneffekt, der darin besteht, dass wir unsere Mitwelt mehr berücksichtigen und mehr Rücksicht auf sie nehmen: Es ist kaum möglich, einen geruhsamen Spaziergang durch den Wald zu genießen, ohne sich ab und zu bücken, um weggeworfenen Abfall aufzuheben und mitzunehmen. Durch die direkte Erfahrung von Wohlbefinden in der heimischen Natur kann dem Fernweh entgegengewirkt werden: So kann der ökologische Fußabdruck, der z. B. gerade durch regelmäßige Flugreisen individuell enorm werden kann, verringert werden, wenn Freizeit- und Urlaubserleben nicht mehr unbedingt exotische Landschaften und Fernreisen assoziiert. Gleichzeitig entwickelt sich dadurch eine neue Wertschätzung für das Lokale, die „everyday landscapes“³⁰ die dadurch auch als Therapeutische Landschaften erfahrbar werden und das eigene Wohlbefinden steigern. In der regelmäßigen Beobachtung der umgebenden Natur wird diese als Mitwelt erlebt und nicht als bloße Umwelt wahrgenommen. Diese Naturwahrnehmung, im Staunen über die Vielgestaltigkeit auch der unmittelbaren Wohnumgebung bildet eine wichtige Säule für ein intensiveres Wertschätzen der Natur.

Literatur

- Abram, David (2015): *Im Bann der sinnlichen Natur. Die Kunst der Wahrnehmung und die Mehr-als-menschliche Welt*, Klein Jasedow
- Bandura, Albert (1977): Self-Efficacy: Toward a Unifying Theory of Behavioral Change, in: *Psychological Review* 84/2, S. 191–215
- Bird, Geoffrey u. a. (2010): Empathic brain responses in insula are modulated by levels of alexithymia but not autism, in: *Brain* 133/5, S. 1515–1525; doi: 10.1093/brain/awq060. Epub PMID: 20371509; PMCID: PMC2859151
- Bucher, Anton A. (2011): *Psychologie der Spiritualität. Ein Handbuch*, Darmstadt
- Charlier, Philippe u. a. (2017): A new definition of health? An open letter of autochthonous peoples and medical anthropologists to the WHO, in: *European Journal of internal medicine* 37, S. 33–37
- Cho, Kyong S. u. a. (2017): Terpenes from forests and human health, in: *Toxicological Research* 33, S. 97–106
- English, Jennifer u. a. (2008): Health, healing and recovery: therapeutic landscapes and the everyday lives of breast cancer survivors, in: *Social Science & Medicine* 67, S. 68–78

³⁰ English u. a. (2008).

- Frankl, Viktor E. (³1978): *Das Leiden am sinnlosen Leben. Psychotherapie für heute*, Freiburg–Basel–Wien
- Hansen, Margaret u. a. (2017): Shinrin-Yoku (Forest Bathing) and nature therapy: a state-of-the-art review, in: *International Journal of Environmental Research and Public Health* 14, S. 851
- Kant, Immanuel (1977): *Kritik der praktischen Vernunft. Grundlegung zur Metaphysik der Sitten*, Frankfurt am Main (Werkausgabe Bd. VII. Herausgegeben von Wilhelm Weischedel)
- Kaplan, Rachel/Kaplan, Stephen (1989): *The experience of nature*, Cambridge
- Kuckartz, Udo (2013): Umweltbewusstsein und Umweltverhalten, in: *Informationen zur politischen Bildung* 287
- Korpela, Kalevi M. (1989): Place-identity as a product of environmental self-regulation, in: *Journal of Environmental Psychology* 9/3, S. 241–256
- Maier, Steven F./Seligman, Martin E.P. (2016): Learned helplessness at fifty: Insights from neuroscience, in: *Psychological review* 123/4, S. 349–367
- Rathmann, Joachim (2020a): Gesundheitsressource Landschaft, in: Jens Soentgen u. a., Hg., *Umwelt und Gesundheit*, Baden-Baden, S. 167–197
- Rathmann, Joachim (2020b): *Therapeutische Landschaften. Landschaft und Gesundheit in interdisziplinärer Perspektive*, Wiesbaden
- Schuh, Angela/Immich, Gisela (2019): *Waldtherapie. Das Potenzial des Waldes für die Gesundheit*, Berlin
- Seligman, Martin E. P. (1975): *Helplessness: On Depression, Development, and Death*, San Francisco
- (1991): *Learned Optimism: How to Change Your Mind and Your Life*, New York
- (2011): *Flourish: A Visionary New Understanding of Happiness and Well-being*, New York
- Spitzer Manfred (2009): Natur und Gemeinschaft. Auswirkungen des Naturerlebens auf prosoziale Motive, in: *Nervenheilkunde* 28, S. 773–777
- (2017): Narzissmus rauf, Empathie runter. Singularisierung in der Nervenheilkunde, in: *Nervenheilkunde* 36/7, S. 550–556
- Otto, Rudolf (2014): *Das Heilige. Über das Irrationale in der Idee des Göttlichen und sein Verhältnis zum Rationalen*. Neuausgabe mit einem Nachwort von Hans Joas, München
- Peterson, Christopher/Seligman, Martin (2004): *Character strengths and virtues: A classification and handbook*, New York–Washington, DC
- Porges, Stephen W. (2010): *Die Polyvagal-Theorie. Neurophysiologische Grundlagen der Therapie. Emotionen, Bindung, Kommunikation und ihre Entstehung*, Paderborn
- Sachverständigenrat für Umweltfragen (1978): *Umweltgutachten (Februar 1978)*, Stuttgart–Mainz
- White, Matthew P. u. a. (2019): Spending at least 120 minutes a week in nature is associated with good health and wellbeing, in: *Scientific Reports* 9, S. 7730, [<https://doi.org/10.1038/s41598-019-44097-3>]

Künstliche Intelligenz im Umweltschutz: Möglichkeiten und Grenzen

Joachim Rathmann

Zusammenfassung

Künstliche Intelligenz (KI) kann durch die Analyse großer Datenmengen, durch das Erkennen von Mustern in den Datensätzen über maschinelles Lernen neues Wissen über Ökosysteme schaffen. Zusätzlich lassen sich dabei Umweltbilanzen erstellen, die als Entscheidungsgrundlage genutzt werden können. Durch die großen Datenmengen können komplexe Rückkopplungseffekte bilanziert werden und Kosten von Entscheidungen transparent gemacht werden. Trotz dieser Chancen von KI für einen ressourcenschonenden Umgang mit der Natur, darf sich aus KI Bilanzierungen kein Entscheidungsautomatismus entwickeln. Für nachhaltiges Umwelts Handeln ist auch eine affektive und emotionale Bindung an die Mitwelt bedeutsam. Dies kann KI nicht leisten, hier ist es Aufgabe der natürlichen Intelligenz ihre Einbettung in einen größeren Naturzusammenhang zu erkennen und daraus einen Lebensstil in umwelttugendethischer Perspektive abzuleiten.

Abstract

Artificial intelligence (AI) can create new knowledge about ecosystems by analyzing large amounts of data, by recognizing patterns in the data sets via machine learning. In addition, this can calculate ecological footprints that can be used as a basis for decision-making. Big data can be used to balance complex feedback mechanisms and make costs of decisions transparent. Despite these opportunities of AI for a resource-saving handling of nature, no decision automatism may develop from AI balances. For sustainable environmental action, an affective and emotional connec-

tion to the environment is also important. This cannot be achieved by AI; here it is the task of natural intelligence to recognize its embedding in a larger natural context and to derive a lifestyle from an environmental virtue ethics perspective.

1. Einleitung

Der Gedanke, mit intelligenter Technik das menschliche Leben zu erleichtern, bewegt die Menschen seit je her. In der Illias sind es die goldenen Dienerinnen, die Hephaistos zu Diensten sind. Er (Hephaistos)

zog einen chiton an, nahm seinen stock und humpelte zur tür
wo zwei dienerinnen herbeieilten um ihren meister zu stützen.
ganz aus gold gegossen, sahen sie wie lebendige mädchen aus;
nicht nur daß sie sprechen und die gliedmaßen bewegen konnten
sie besaßen auch verstand und hatten dazu von den unsterblichen
die vielseitigsten fertigkeiten gelernt (Homer XVIII, V. 415ff.)¹.

Aspekte, die hierbei schon angesprochen werden sind außerordentlich aktuell. Pflegeroboter beispielsweise sollen ältere Menschen im Alltag unterstützen und über maschinelles Lernen sich an neue Situationen anpassen können. Der Einsatz von Robotern soll dabei das Pflegepersonal durch Tätigkeiten wie das Verteilen von Essen oder Medikamenten oder das Leeren von Mülleimern entlasten und gleichzeitig ältere oder behinderte Menschen länger in der Selbstständigkeit unterstützen. Die Akzeptanz dieser Roboter soll dabei durch eine gewisse Menschenähnlichkeit und durch Eigenschaften wie Lernfähigkeit und Autonomie erhöht werden². Zusätzlich kann eine KI-gestützte Spracherfassung in der Pflege Entlastung bei Routine-tätigkeiten wie Dokumentation und Verwaltung bewirken. Technisch basiert dies auf Formen des maschinellen Lernens, die als ein zentraler Teilbereich der Künstlichen Intelligenz (KI) angesehen werden. Doch bleibt in diesem Kontext der eigentliche Begriff der „Intelligenz“ oft unscharf, kann jedoch generell als das Ausmaß der Problemlösungsfähigkeit künstlicher Systeme angesehen werden. Zahlreiche

¹ Schrott (2011), S. 391.

² Bendel (2018).

Systeme der KI sind inzwischen fest etablierter Bestandteil der Lebenswirklichkeit vieler Menschen; sei es im Erlernen von Vorlieben bei Musik oder Filmen oder im Kaufverhalten. Ansätze einer schwachen KI, auf welche hier zurückgegriffen wird, dienen dazu, konkrete Anwendungsprobleme lösen. Solche Ansätze versuchen also gar nicht, sämtliche Eigenschaften menschlicher Intelligenz abzubilden, sondern fokussieren auf einen Teilbereich, der durch schnelle Rechenoperationen zu bewältigen ist. Unterschiedliche Verfahren des maschinellen Lernens bilden dabei wichtige Ansätze einer schwachen KI. Aus großen Datenmengen werden statistische Abhängigkeiten und Muster ermittelt, welche sich für Vorhersage- oder Klassifikationszwecke anwenden lassen. Die Güte dieser Anwendungen hängt dabei von der Menge und Qualität der Eingangsdaten ab.

Im Folgenden soll die Frage adressiert werden ob bzw. inwieweit Künstliche Intelligenz helfen kann, die sich zunehmend verstärkende globale ökologische Krise zu meistern. Denn das Wissen um die Zerstörung von Natur und Umwelt ist seit vielen Jahrzehnten vorhanden und der Naturschutzgedanke hat eine lange Geschichte. In Deutschland gilt der Drachenfels bei Königswinter als erstes Naturschutzgebiet, das 1836 eingerichtet wurde. Damit sollte jedoch eher ein romantisch aufgeladenes Symbol erhalten werden und nicht primär eine unberührte Natur. D.h. der Naturschutz dient seit je her auch dem Schutz von Kulturlandschaften. Der weltweit erste Nationalpark (Yellowstone 1872, gefolgt vom Yosemite 1890) führt dann zu einem verstärkten Bewusstsein, Gebiete als Lebensraum für Tiere und Pflanzen zu schützen³. Mit „Pfisters Mühle“ (Wilhelm Raabe 1884) erschien im nämlichen Zeitraum der erste deutsche Umweltroman, ein Zeugnis gründerzeitlicher Gewässerbelastung durch Zuckerfabriken. Damit ist in den westlichen Ländern seit weit über 100 Jahren ein Bewusstsein dafür vorhanden, Natur und Landschaften zu schützen. Ebenso eindrücklich wie aktuell schildert Ludwig Klages 1914 die Situation: „Eine Verwüstungsorgie ohnegleichen hat die Menschheit ergriffen, die ‚Zivilisation‘ trägt die Züge entfesselter Mordsucht, und die Fülle der Erde verdorrt vor ihrem giftigen Anhauch. So also sehen die Früchte des ‚Fortschritts‘ aus!“⁴. Eingerahmt in diese Zerstörung ist auch das Aufkommen von Pandemien, denn das Eingreifen in bislang kaum berührte Ökosysteme kann neue Übertragungswege für Zoonosen eröffnen und Pandemien einleiten. Auch dieser Gedanke findet sich bereits bei

³ Zur Geschichte des Naturschutzes in Deutschland, vgl. Stammel / Cyffka (2015), Kap. 1.2.

⁴ Klages (1974) S. 619.

Klages: „und so geht es fort bis zu den schlimmeren Rückschlägen der verwundeten Natur exotischer Länder in Gestalt jener furchtbaren Seuchen, die sich an die Ferse des ‚zivilisierten‘ Europäers heften“⁵. Über 100 Jahre später sind es nicht mehr nur die „zivilisierten“ Europäer, die mit den Folgen menschlicher Eingriffe in wenig berührte Ökosysteme zu kämpfen haben. Dass Menschen mit der Zerstörung von Natur letztlich sich selbst schaden, ist auch in der jüngeren Umweltgeschichte mit dem Erscheinen des Sachbuches „Silent Spring“ von Rachel Carson 1962 auf große Resonanz in der Öffentlichkeit gestoßen. Das Wissen um die Dringlichkeit, global einen effektiven Klima- und Naturschutz durchzusetzen, ist damit seit Jahrzehnten einer breiten Öffentlichkeit ebenso wie Entscheidungsträgern zugänglich. Es mangelt folglich weniger am Umwelt-Wissen oder Umwelt-Bewusstsein, es mangelt am Umwelt-Sein, am Umwelt-Handeln.

Dass auf Umweltkrisen nicht unmittelbar, affektiv reagiert wird, liegt einerseits darin begründet, dass die Schäden für den Menschen oft raum-zeitlich asynchron auftreten sowie daran, dass Menschen in ihrer Handlung einen „Hiatus“ zwischen der unmittelbaren Bedürfnisbefriedigung und einer notwendigen Alltagshandlung aufbauen. In dieser „Indirektheit der Lebensfristung“⁶ mag ein Grund für die massive Diskrepanz zwischen Wissen und Handeln im regionalen und globalen Umweltdiskurs liegen. Ergänzend lässt sich der Aspekt der Abwehr von Angst anführen als ein Verdrängungsmechanismus vor apokalyptischen Szenarien. Bei vorhandenem Wissen kommt zusätzlich eine selektive Unaufmerksamkeit und Selbstbetäubung hinzu⁷. Darin mag eine weitere Ursache dafür liegen, dass wir trotz des immensen Wissens um den Zustand globaler Ökosysteme uns eben doch nicht angemessen verhalten⁸. Insgesamt bleibt ökologisches Wissen und Handeln nur lose aneinandergekoppelt; über kognitive Dissonanzen gilt dies auch für besondere umweltbewusste Menschen. Kognitive Dissonanzen entstehen, wenn die Einstellung, Meinung oder Norm nicht zum eigentlichen Handeln passt⁹. Menschen streben danach, solche Spannungszustände abzubauen. Dazu werden im ökologischen Diskurs gerne Argumente gesucht, das eigene Handeln zu begründen, beispielsweise über Sachzwänge,

⁵ Ebd., S. 617.

⁶ Gehlen (1978), S. 335.

⁷ Dreitzel (1992), 228ff.

⁸ Ausführlich zum „Motivationsproblem“ im Umwelthandeln: Baumgartner (2004).

⁹ Festinger (1962).

institutionelle Anreize und weitere Notwendigkeiten. Am Beispiel deutscher Nachhaltigkeitsforscher lässt sich zeigen, wie diese ihren anwachsenden ökologischen Fußabdruck durch Flug(dienst)reisen mit solchen Argumenten rechtfertigen¹⁰. Jedoch: der Ausstoß an Treibhausgasen bleibt davon unberührt. Eine weitere Erklärung für solches Verhalten könnte sich mit moralischem Lizensieren beschreiben lassen¹¹. Gerade umweltbewusste Menschen, da sie ja für die Sache des Guten eintreten, buchen dann gleichsam von einem imaginären Umweltkonto beispielsweise eine Flugreise, für die sich sicherlich gute Sachzwänge anführen lassen, ab. Mentale *rebound*-Effekte können dann zu einem verstärkten Ressourcenverbrauch führen. Damit werden Effekte beschrieben, die zur Folge haben, dass das ursprüngliche Einsparpotenzial, etwa auf Grund von Effizienzsteigerungen nicht oder nur teilweise verwirklicht wird. Dies kann zur Konsequenz haben, dass in der ökologischen Gesamtbilanz dadurch die Haltung, für eine ökologisch gute Sache einzustehen, dann das insgesamt nachhaltige Handeln gleichsam ersetzt. Kurz: es mangelt nicht an gutem Willen: „Guter Wille ist erfreulicherweise reichlich vorhanden; er demonstriert überall“, es mangelt nicht an „Gesinnung“¹². Doch der lebensweltliche Vollzug des Bewussten zeigt sich weniger in der Wahrhaftigkeit der Gesinnung als vielmehr in der Tatkraft. Doch trotz allem ist es selbst im Alltag oftmals gar nicht so einfach festzustellen, welche Entscheidung den geringsten Ressourcenverbrauch nach sich zieht. Hierbei könnte KI jedoch eine wertvolle Entscheidungshilfe bereitstellen.

2. Das Paradoxon des Umweltwissens

Das „Paradoxon im Umweltwissen“¹³ beschreibt ein Phänomen unterschiedlicher raum-zeitlicher Skalen. Denn auf einer globalen Ebene sind die Erfordernisse, um einen wirksamen Naturschutz im umfassenden Sinne zu erreichen, seit vielen Jahrzehnten klar zu benennen: beispielsweise die Reduktion von Treibhausgasemissionen, des Flächenverbrauchs, der Zerschneidung von Habitaten, großflächiger Rodungen, einer Intensivierung der Landnutzung oder einer Überfischung in den

¹⁰ Schrems/Upham (2020).

¹¹ Halla, Lewis, Jr. u. a. (2018).

¹² Lübke (2019), S. 96.

¹³ Verf. (2020b).

Ozeanen. Trotz dieser Wissensbestände, zeigt sich in der „großen Beschleunigung“ („the great acceleration“¹⁴), dass die wesentlichen Indikatoren für den Zustand globaler Ökosysteme, trotz regionaler (und bezogen auf die Ozonschicht auch globaler) Verbesserungen, weiterhin beschleunigte Trends in eine negative Richtung zeigen. Offenbar ist das Wissen um die ökologische Krise nur unzureichend handlungsrelevant. Ein Grund dafür lässt sich damit benennen, dass es im individuellen Verhalten auf einer lokalen Ebene oft gar nicht so klar ist, was bezogen auf komplexe Folgewirkungen einer Entscheidung die wirklich ökologischere Alternative ist. Denn im „Paradoxon des Umweltwissens“ zeigt sich, dass sich (nicht nur) auf einer individuellen Ebene vermeintlich ökologisch-nachhaltige Entscheidungen als solche mit komplexen negativen Auswirkungen zeigen können. Die in Deutschland gekaufte Biokarotte aus Israel mag „bio“ sein, „öko“ ist sie sicherlich nicht. Ist aber das regionale Produkt generell ökologisch nachhaltiger als eines aus entfernteren Regionen, in denen effizienter mit geringerem Ressourceneinsatz angebaut werden kann?

Zusätzlich lässt sich eine monetäre Perspektive ergänzen: Eine Bambuszahnbürste mag nachhaltiger sein als eine Plastikzahnbürste, welche nur ein Drittel des Preises kostet. Jedoch könnte, in der Perspektive des effektiven Altruismus, das dabei gesparte Geld für Umweltschutzmaßnahmen eingesetzt werden und dadurch einen insgesamt größeren ökologischen Nutzen stiften.

„Greenwashing“ bezeichnet in negativer Konnotation das Hervorheben ökologischer Vorteile von Produkten oder Prozessen, ohne dass im Gesamt der Betrachtung sämtlicher Wechselwirkungen dafür eine Grundlage vorhanden wäre. Dies erfolgt zumeist unter Hervorhebung selektiver Aspekte. Beispielsweise sind bei einem T-Shirt aus Bio-Baumwolle in einer Gesamtbilanz u. a. der hohe Wasserverbrauch für Baumwolle (in zumeist trockenen Regionen), der Flächenbedarf und damit die Flächenkonkurrenz, der Transport oder der Einsatz von stofffärbenden Substanzen zu berücksichtigen.

Als weiteres Beispiel könnte das Recycling von Papier und Kartonagen angeführt werden. Dieses Verhalten kann eine gewisse Umweltentlastung bewirken und noch viel stärker beim Handelnden das Gefühl evozieren, ein guter Umweltschützer zu sein. Jedoch ist bei einem stetig ansteigenden Verpackungsmüll durch den weiter zunehmenden Onlinehandel das Recycling ein geringerer Teil der Lösung, ein We-

¹⁴ Steffen u. a. (2015).

niger im Verpackungseinsatz und der Bestellmenge wäre hier zielführender hinsichtlich einer wirklich nennenswerten Umweltentlastung¹⁵.

So laufen Bio-Etikette und Recycling Gefahr, größeren Schaden anzurichten bei gleichzeitigem Steigern des Umweltbewusstseins der Menschen. Sie erwecken die Illusion des Guten, feiern einen Triumph der Gesinnung und verkennen die Komplexität von Zusammenhängen, so dass sie letztlich stärkere negative Auswirkungen haben können, als gemeinhin bewusst ist. Das Gutgemeinte ist nicht deckungsgleich mit dem Guten. Denn es gilt für den Ressourceneinsatz bei vielen Produkten und Prozessen, dass die Wechselwirkungen auch in unterschiedlicher räumlicher und zeitlicher Ausprägung so komplex sind, dass die schnelle Entscheidung für das vermeintlich ökologischere Produkt falsch liegen kann. Die sich daraus ergebende „Absichtslosigkeit, der rasch anwachsenden Zivilisationslasten“¹⁶ sollte vor dem weit um sich greifenden Moralismus in ökologischen Fragen warnen. Denn zu gerne werden die Zivilisationslasten dem Kapitalismus, dem „System“ und den Großkonzernen angelastet.

Beim Abwägen ökologischer Konsequenzen von Handlungen und Kaufentscheidungen könnte man leicht in die Rolle von Buridans Esel geraten, der zwischen zwei gleich weit entfernten ebenso großen Heuhaufen verhungert, weil er sich nicht entscheiden kann, welchem er sich zuwenden soll. Analog könnte ein detailliertes Abwägen in Umweltentscheidungen zu einer Deadlock-Situation führen, in der sich beide Alternativen gegenseitig blockieren und eine Situation ausweglos erscheinen mag. KI könnte hierbei ansetzen und Entscheidungsgrundlagen kraft der Berechnung großer Datenmengen schaffen und zur Transparenz der wahren Kosten und des tatsächlichen Nutzens beitragen.

3. Hoffen auf KI?

Die rasche Bearbeitung großer Datenmengen durch KI und das Erkennen von Mustern in den Datensätzen kann neues Wissen über Ökosysteme schaffen und den Um-

¹⁵ Lübke konstatiert, dass das Verständnis für die eigenen Lebensbedingungen abnimmt und wir daher zunehmend auf die „Könnerschaft“ von Experten angewiesen sind, dies kann nur auf Vertrauen basieren; vgl. Lübke (2019), S. 117. Die Notwendigkeit, Schadenebenenfolgen individuellen Handelns zu berücksichtigen und Folgen abzuschätzen steigt daher stark an.

¹⁶ Ebd., S. 77.

gang mit ihnen optimieren. Dadurch kann nachhaltiges Umweltverhalten durch KI simuliert werden und die reale Nutzung von Ökosystemen grundsätzlich ressourcenschonender und nachhaltiger gestaltet werden. In der Landwirtschaft beispielsweise finden sich unterschiedliche Anwendungsfelder für KI: landwirtschaftliche Prozesse lassen sich dadurch in Echtzeit ortsbezogen bedarfsgerecht steuern. Eine ortsdifferenzierte und zielgerichtete Bewirtschaftung landwirtschaftlicher Nutzflächen wird als Precision Farming bezeichnet und ist ein Teil einer Digitalisierung der Landwirtschaft. Dies erfolgt auch unter Einsatz von Drohnen, um genaue Daten zu sammeln und hochaufgelöste Bilder zu erstellen, die bei der Überwachung der Kulturen helfen und gleichzeitig dazu beitragen, den Ressourceneinsatz zu optimieren und damit eine Umweltentlastung zu ermöglichen¹⁷. Denn Precision Farming verwendet KI, um genaue und kontrollierte Techniken zu entwickeln, die helfen, Orientierung und Verständnis für das Wasser- und Nährstoffmanagement, optimale Ernte- und Pflanzzeiten sowie den Zeitpunkt der Fruchtfolge zu liefern.

Eine weitere Umweltentlastung im Agrarsektor könnte sich auch durch das Vertical Farming ergeben, dabei werden in geschlossenen Systemen in Indoor-Farmen Gemüse und Salate angebaut. Eine hohe Produktivität ist dadurch gegeben, dass die Systeme platzsparend über mehrere Etagen (vertikal) auf künstlichen Nährböden bzw. in Nährlösung wachsen können. Die Nähe zu den Verbrauchern ist ein weiterer Vorteil. KI kann dabei den Einsatz von Wasser, Nährstoffen, Licht, Energie oder Luftfeuchtigkeit exakt steuern und die Pflanzen optimal ohne Einsatz von Pflanzenschutzmitteln versorgen¹⁸.

Eine weitere Chance zu einer Verbesserung des ökologischen Zustandes liegt darin, dass KI bei der Berechnung von Umweltbelastungen über Ökobilanzen, Klimabilanzen oder den ökologischen Fußabdruck eingesetzt werden kann. Denn je größer die eingehenden Datenmengen und je mehr Wechselwirkungen dabei Berücksichtigung finden können, umso präziser können solche Berechnungen ausfallen. KI hat das Potenzial, diese Datenvolumina zu bewältigen und durch Prozesse des Selbstlernens die Übertragbarkeit von Produktbewertungen vorzunehmen. Dadurch kann der jeweilige „Umweltverbrauch“ von Produkten aber auch Dienstleistungen quantifiziert werden und als Entscheidungsgrundlage vorgehalten werden. Externe Kosten können ebenso transparent dargestellt werden. Kunden könnten

¹⁷ Finger u. a. (2019).

¹⁸ Al-Chalabi (2015).

dann beim Kauf von Produkten unmittelbar auf der Grundlage umfassender Informationen eine ökologisch nachhaltige Wahl treffen. Multi-kriterielle Entscheidungsanalysen (multi-criteria decision analysis, MCDA) finden daher zunehmend Verbreitung bei vielen Aspekten aus dem Energie- und Umweltsektor¹⁹. Denn die wesentlichen Entscheidungsprobleme entstehen, wenn mehrere Ziele (multiple Kriterien) verfolgt werden und sich die Entscheidungen in einem komplexen Kontext abspielen. Häufig sind die vorhandenen Informationen und Ziele konfliktärer Natur, wenn es gilt ökonomische, soziale Belange mit Ansprüchen von Arten- und Klimaschutz auszutarieren; hinzukommen substantielle Konsequenzen und langfristige Auswirkungen in unterschiedlicher räumlicher Ausprägung. Daher lassen sich Fehlentscheidungen dann nicht mehr so leicht revidieren. Komplexe Entscheidungen sind nicht mehr trivial und können durch die pure Rechenkapazität der KI sinnvoll Argumente für Entscheidungen zutragen. Denn für die Entscheidungsträger werden die Situationen durch eine multi-kriterielle Entscheidungsanalyse formalisiert, in dem Informationen soweit organisiert werden, dass der Entscheidungsträger mit dem Gefühl die wesentlichen Kriterien berücksichtigt zu haben, seinen Teil zu einem verbesserten Entscheidungsprozess beitragen kann²⁰. Das Ziel ist es dabei, eine technische Unterstützung bei komplexer Problemlage zu bieten, konsistente, nachvollziehbare und vernünftige Entscheidungen zu treffen bzw. Kompromissverhandlungen zu unterstützen basierend auf der Möglichkeit durch Auswahl, Vergleich und Einordnen verschiedener Attribute mehrerer Alternativen auf flexible Weise gegeneinander abzuwägen.

4. Grenzen und Risiken der KI

KI-gestützte Entscheidungsanalysen in Umweltfragen können die grundsätzliche Diskrepanz von Umweltwissen und -handeln kaum überbrücken. Jede Entscheidungshilfe bedarf der Umsetzung durch den Entscheidungsträger. Doch zeigt sich im Umweltdiskurs der vergangenen Dekaden, dass das Wissen, wie die ökologische Situation auf globaler aber auch lokaler Skala verbessert werden kann, durchaus vorhanden ist. Faktenwissen alleine ist ein notwendiges jedoch kein hinreichendes

¹⁹ Diakoulaki u. a. (2005).

²⁰ Belton/Stewart (2002).

Vehikel, Weltwendung herbeizuführen. Faktenwissen, auch dieses einer KI-generierten Entscheidungshilfe berührt den Menschen kaum lebensweltlich, es bleibt ihm äußerlich²¹. KI kann das „Was“ in ökologischen Belangen berechnen, für die Frage des „Wozu“ ist jedoch die natürliche Intelligenz von Nöten²².

Hinzu kommt, dass Information oft in ein gewisses Deutungsraaster (*framing*) eingebettet ist; für die KI besteht zunächst kein Unterschied darin, ob ein Glas halbvoll oder halbleer ist, für eine menschliche Entscheidung gleichwohl. Rein logisches Abwägen ist blind für Intuition und die individuelle oder sozio-kulturelle Einbettung von Entscheidungen. Schließlich bedarf die rein instrumentelle Vernunft der KI einer Ergänzung, ansonsten besteht die Gefahr einer „technische[n] Perfektion bei vollständigem Ausfall moralischer Reflexion“²³. Denn das menschliche Gewissen in seiner notwendigen Abwägung bewahrt davor, dass die Urteile der KI als Richtersprüche zu Scharfrichtersprüchen werden und der Mensch schließlich zum Sklaven der digitalisierten Welt wird. Denn die Leibhaftigkeit des Ichs als Person verhindert, dass die Welt nur aus einer Zuschauerperspektive wahrgenommen wird. KI jedoch ist Ausdruck eines Szientismus in der Tradition von Francis Bacon. Dabei besteht die Gefahr, dass letztlich die „schlechteste aller möglichen Welten“²⁴ konstruiert wird. Denn „Gegenständlichkeit wird dann gleichgesetzt mit Verfügbarkeit – eine Gleichung, die auf die Abschaffung von Gegenstand und Gegenständlichkeit hinausläuft [...]. Das Gegenteil ist wahr. Der total unverfügbare Gegenstand ist am meisten Gegenstand – die (als menschlich oder übermenschlich verstandene) PERSON. In ihr, und nur in ihr, ist ein Höchstmaß der Tiefe verwirklicht“²⁵. Damit entzieht sich die natürliche Intelligenz auch der künstlichen von der sie sich mannigfaltig unterscheidet. KI bildet allenfalls eine „*Simulation* von eng umgrenzten Bereichen menschlicher Intelligenz“²⁶. Denn wesentliche Aspekte wie Leben, Bewusstsein oder Perspektivenübernahme lassen sich durch Algorithmen nicht erzeugen²⁷. Leben findet im Leben selbst statt und kann nicht durch eine Model-

²¹ Vgl. Verf. (2020b).

²² Vgl. Kuhn (1981), S. 246.

²³ Lübke (2019), S. 9.

²⁴ Kuhn (1981), S. 352.

²⁵ Ebd., S. 353.

²⁶ Fuchs (2020), S. 45.

²⁷ Vgl. Tabelle 1 bei Fuchs (2020), S. 59f.

lierung substituiert werden, Leben heißt immer auch affektiv und emotional zu Anderem in Beziehung treten. Denn: „die grenzenlose Versachlichung der Menschen unserer Tage frisst allmählich auch die Kräfte weg, die zur Aufrechterhaltung einer bloßen Sachkultur und bloß technischer Betriebe erforderlich sind, z. B. Phantasie, Schöpfungstum, Hinlauschen auf die Lebensquellen, die in der Tiefe rauschen. Warum erquickt uns der Umgang mit der Natur? Weil wir da einmal mit uns allein sind und also auch mit uns selbst Umgang pflegen können“²⁸. Diese Bindungsfähigkeit könnte sich aber als ein zentraler Aspekt im Überwinden der ökologischen Krise herausstellen. Doch hier zeigen sich die Grenzen der KI, denn Empathie hat eine leibliche Komponente, die dadurch nicht abgebildet werden kann, trotz einer „fiktionalen Empathie“²⁹, die sich auch gegenüber Computerfiguren oder Robotern einstellen kann. Eine sinnentleere, in die virtuelle Welt abgewanderte „Als-ob“-Empathie verliert jedoch an Tiefe und Verbindlichkeit.

4.1. Empathie

Die zunehmende Präsenz digitaler Medien, Zeichensysteme und Fiktionen könnte in den vergangenen Jahren zu einem Rückgang von Perspektivenübernahmen und primärer Empathie sowie des psychischen Wohlbefindens geführt haben. Eine vielbeachtete Metastudie³⁰, basierend auf Daten von fast 14.000 Studenten aus 72 Studien der Jahre 1979 bis 2009 zusammengetragen, stellt über diesen Zeitraum einen Rückgang von Empathie heraus. Besonders nach dem Jahr 2000 tritt dieser Rückgang offenkundig in Erscheinung. Als Index wird der *Interpersonal Reactivity Index* (IRI) herangezogen, welcher anzeigt inwiefern sich jemand in einen Anderen oder in Figuren in Filmen oder Büchern hineinversetzen kann, um jenen und diese zu verstehen. Die Bereitschaft zur Perspektive-Übernahme ist in den letzten Jahren des Untersuchungszeitraums gesunken während die Werte für die Fantasie konstant geblieben sind. Diese Abnahme korreliert positiv mit der gängigen Zeitdiagnose „Narzissmus“³¹. In dieser Studie wurde letztlich die Einstellung zur Empathie er-

²⁸ Spranger (1962), S. 25.

²⁹ Fuchs (2020), S. 125ff.

³⁰ Konath u. a. (2011).

³¹ Vgl. Verf. (2020b), S. 115.

fragt und nicht das wirkliche (empathische) Verhalten selbst beobachtet. Sicherlich lässt sich auch kritisch fragen, ob bei einem so langen Untersuchungszeitraum beim selben Fragebogen zwischen 1979 und 2009 nicht Bedeutungsverschiebungen und unterschiedliche Assoziationen bei den Befragten auftreten. Aber trotz aller grundsätzlicher methodischer Kritik, könnte es sein, dass hier ein altes kulturpessimistisches Lamento Bestätigung sucht, denn „immerhin wurde bereits vor gut 200 Jahren behauptet, dass die neuen Medien die Jugend verderben und zum Narzissmus führen würden. Die neuen Medien waren damals die Romane, von denen wir uns jetzt wünschen würden, dass die Jugend sie mehr liest“.³² Dennoch: ein Weniger an Empathie und eine Mehr an Narzissmus scheint sich besonders in sozialen Medien zu offenbaren und dabei zeigt sich zusätzlich ein negativer gesundheitlicher Effekt: Das psychische Wohlbefinden sinkt bei Jugendlichen, wenn viel Zeit vor dem Bildschirm bzw. Smartphone (soziale Medien, Internet, Spiele, ...) verbracht wird verglichen mit Menschen, die verstärkt Aktivitäten jenseits eines Bildschirms ausüben (direkte persönliche Kontakte, Sport, kirchliche Aktivitäten o. ä.). Dies konnte in einer landesweiten Erhebung aus den USA über einen Zeitraum von 15 Jahre zwischen 1991 und 2016 gezeigt werden³³.

Eine Bewertung der (vermeintlichen) Empathieabnahme hängt jedoch davon ab, wie der stark positiv konnotierte Begriff³⁴ inhaltlich zu füllen ist. Empathie als Einfühlungsvermögen ist etwas anderes als Mitgefühl mit Fürsorglichkeit. Denn Bloom³⁵ argumentiert dagegen und zeigt, dass Empathie als bloßes Einfühlungsvermögen schreckliche Situationen rechtfertigen kann. Konflikte können durch ein Einfühlen in bestimmte Gruppen verstärkt werden. Eine bloße Perspektive-Über-

³² Breithaupt (2019), S. 71. Seit dem Verlassen des Gartens Eden bleibt das „O tempora o mores!“ die Begleitmelodie des Anthropozäns. Moderner formuliert dies Szyborska (2014), S. 65 in dem Gedicht „Nichtlesen“:

„Wir leben länger
aber ungenauer
und in kürzeren Sätzen.
Wir reisen schneller, öfter, weiter.
und statt Erinnerungen bringen wir Fotos mit.“

³³ Twenge et al. (2018).

³⁴ Für de Waal (2011) ist die Fähigkeit, sich mit anderen verbunden zu fühlen, das Bindemittel, das Menschen und Völker positiv verbindet. Für ihn ist „Empathie für ‚andere Völker‘ [...] der Rohstoff, den die Welt noch dringender braucht als Öl“ (S. 263).

³⁵ Bloom (2018).

nahme, ein Eindenken in Andere ist auch bei Gewalttätern möglich, denn Einfühlungsvermögen beschreibt die Fähigkeit und die Tendenz dazu, die Gefühle, von denen man glaubt, das Gegenüber fühle sie, selbst auch zu empfinden³⁶.

Die Doppelseitigkeit von Perspektive-Übernahmen hat bereits Lipps³⁷ im Begriff der „Einfühlung“ herausgearbeitet. Er unterscheidet die „lustgefärbte“ positive von der „unlustgefärbten“ negativen Einfühlung. Das „positive Einfühlen“ ist „jenes Aufnehmen des in mich Eindringenden, oder es ist das Einswerden des erfassenden Ich, wie es an sich ist, mit dem, was sich in dasselbe eindringt“³⁸. Die „negative Einfühlung“ hingegen wird als das beschrieben, gegen dessen Eindringen sich „Widerspruch“ erhebt. Es weist sich als mit sich selbst „unverträglich“ zurück³⁹. Implizit findet sich diese Ambivalenz auch bei Bloom, indem er von der Empathie und deren negativen Seiten Aspekte wie „kindness“ (Freundlichkeit) und „compassion“ (Mitgefühl) unterscheidet, die er ausdrücklich positiv würdigt ebenso wie die positiven Seiten, die Empathie ja auch zeigt. Doch Empathie kann nach Bloom eben auch zu Gleichgültigkeit oder gar Grausamkeit motivieren, denn Empathie liegt eine gewisse Kurzfristigkeit zu Grunde, da sie den Fokus auf ein bestimmtes Gegenüber setzt; dabei besteht die Gefahr, längerfristige Konsequenzen und das Leid derer, die nicht das aktuelle Gegenüber sind, zu übersehen. Die Nächstenliebe läuft Gefahr, die Fernstenliebe zu überblenden⁴⁰. Nietzsches Zarathustra empfiehlt: „zur Nächstenliebe rathe ich euch nicht: ich rathe euch zur Fernsten-Liebe“⁴¹. Nicolai Hartmann

³⁶ Breithaupt (2019) zeigt an zahlreichen Beispielen die „dunklen Seiten der Empathie“. Dies soll den Blick dafür schärfen, dass eine zentrale Eigenschaft menschlichen Lebens, Empathie zu entwickeln auch negative Folgen nach sich ziehen kann.

³⁷ Lipps (1907).

³⁸ Ebd., S. 236.

³⁹ Ebd.

⁴⁰ Bloom bringt das auf die kurze Gleichung: „Self + Close People + Strangers = 100 %“ (S. 162). Welche Anteile (Zeit, Geld, Engagement, Emotionen, ...) werden in welchen Teilbereich wie stark investiert? Wobei klar ist, dass es Ressourcen gibt, die sich beim Einsatz vermindern (z. B. Geld) und solche, die sich beim Einsatz sogar vergrößern (z. B. Liebe, Zuneigung). Dazu bemerkt Batthyány (2017), S. 89: „Einen Menschen getröstet oder ermutigt zu haben, bedeutet nicht, dass uns die Worte des Trostes oder der Ermutigung irgendwann ausgehen für den Nächsten, der Trost oder Ermutigung bedarf“. Liebe und Zuneigung sprengen daher das Modell von Ressourcen, indem durch das Geben das Subjekt an Reichtum gewinnt und daran verliert sofern die Möglichkeit des Gebens nicht realisiert wird.

⁴¹ Nietzsche (1999), S. 79.

sieht bei Nietzsche trotz aller Überspitzung anerkennend das „positiv Geschauete[...]“⁴² und führt die Fernstenliebe als Liebe aus, „die keine Gegenliebe kennt, die nur ausstrahlt“⁴³. Freilich setzt die Fernstenliebe beim Nächsten an, sieht diesen aber gewissermaßen als Mittel für einen höheren (zukünftigen) Zweck. Oftmals ist die Fernstenliebe anstrengungslos zu vollziehen. Eine Petition für Flüchtlinge oder gegen die Abholzung des Regenwaldes zu unterschreiben verschafft Selbstbestätigung für das Gute einzutreten, den Müll am Straßenrand aufzuheben ist hingegen vergleichsweise unbequem. Daher muss der Ausgangspunkt des Handelns zunächst in der Einfühlung der unmittelbaren Umgebung liegen. Die Bindung des Ich im Du ist nicht einfach eine Projektion des Eigenen im Anderen, es ist als Liebeserfahrung eine Zusicherung des Vorranges dem Du gegenüber in Treue; andernfalls bliebe die Verantwortlichkeit „ein freischwebendes Sollen“⁴⁴.

Auch wenn moralische Entscheidungen wesentlich durch Empathie geprägt sind, ist es wichtig zu erkennen, dass sich dabei auch negative Konsequenzen ergeben können. Mitgefühl (*compassion*) ist daher ein geeigneterer Weg, zum Besseren beizutragen, da es nicht einfach die Gefühle des Gegenübers nachvollzieht, gleichsam mit ihm empfindet, sondern durch Anteilnahme die Motivation erwächst, das Wohlergehen des Gegenübers zu fördern indem für den Gegenüber empfunden wird. Dies kommt auch dem Empathie-Begriff nahe, den Goleman im Kontext einer „ökologischen Intelligenz“ anführt. Diese umfasst einerseits das Wissen um ökologische Zusammenhänge und schließt andererseits die Einsicht ein, „aus Erfahrungen zu lernen und sinnvoll zu handeln“⁴⁵. Das ist für ihn an eine Form von Empathie gebunden, die alles umfasst „was lebt“⁴⁶. Dies impliziert ein Mitleiden, wenn Ökosysteme „leiden“ und daraus ein Handeln abzuleiten, welches dieses Leiden zu mindern trachtet. In einer tiefenökologischen Perspektive weitet sich die Perspektive auch auf Unbelebtes: „denken wie ein Berg“, Aldo Leopolds Diktums, verdeutlicht, dass der Schutzgedanke auch über die belebte Welt hinausgeht⁴⁷. Auch für Berry⁴⁸ ist der Gedanke zentral, dass menschliche

⁴² Hartmann (1962), S. 484.

⁴³ Ebd., S. 490.

⁴⁴ Vetter (1949), S. 224.

⁴⁵ Goleman (2009), S. 48.

⁴⁶ Ebd., S. 50.

⁴⁷ Vgl. Devall (1997).

⁴⁸ Berry (2002).

Körper eng mit der umgebenden Natur verwoben sind und erst der Kontakt mit der „Wildnis“ Erfahrungen bringt, die „uns demütig werden lässt und die uns belebt, die schmerzlich und freudvoll zugleich ist“⁴⁹. Abram eröffnet die Perspektive, dass

das wahrnehmende Wesen und das wahrgenommene Wesen aus ein und demselben Stoff sind, dass Wahrnehmendes und Wahrgenommenes voneinander abhängen und gewissermaßen sogar umkehrbare Aspekte eines gemeinsamen, belebten Elementes, oder Fleisches, sind, das zu gleich empfindbar und empfindend ist⁵⁰.

Diese Reziprozität des Sinnlichen erstreckt sich für Abram unmittelbar auch auf nicht-menschliches Leben, das sich in einem Kontinuum bis in die Landschaft erstreckt. Für ihn schließt dies explizit auch die Fernstenliebe ein:

Denn wenn ich meine Umgebung als sinnlich empfindend, als achtsam und wachsam erfahre, so muss ich selbst ebenfalls Achtsamkeit und Respekt walten lassen – auch dann, wenn ich weit entfernt von anderen Menschen bin –, um das wachsame Land nicht zu kränken⁵¹.

Hier schließt sich unmittelbar die Frage nach dem guten Leben an. Aus einer vertieften (nicht unbedingt einer tiefenökologischen) ökologischen Perspektive erhellt sich von selbst, dass der Sinn eines guten Lebens sich nur in einem maßvollen und ehrfürchtigen Umgang mit der Mitwelt erschließt.

4.2. Mitgefühl als tugendethisches Potential

Angeichts globaler ökologischer Herausforderungen stellt sich die Frage nach der Umsetzbarkeit ethischen Handelns in neuer Dringlichkeit. In der Perspektive einer Normethik, findet eine etablierte Norm ihre Anwendung in einem speziellen Fall. Jedoch zeigt sich in der Komplexität ökologisch-sozialer Systeme mit immer neuen

⁴⁹ Ebd., S. 24.

⁵⁰ Abram (2015), S. 85.

⁵¹ Ebd., S. 87.

Rückkopplungen und Rebound-Effekten, die Notwendigkeit, Adaptionmöglichkeiten zu maximieren. Eine tugendethische Perspektive eröffnet die Möglichkeit, Persönlichkeitsmerkmale zu stärken, welche helfen in komplexen Situationen allen Belangen gerecht zu werden.

Empathie ist aufgrund der Ambivalenz der Perspektive-Übernahmen zunächst nicht moralisch. Doch lässt Empathie als Mitfühlen und als Liebe weitere Dimensionen stärken, denn die Übernahme der Perspektive Anderer bereichert das eigene Empfinden und die eigene Wahrnehmung um neue Perspektiven. Ergänzt um das Mitfühlen mit Anderen wird die sinnliche Dimension der Wahrnehmung vergrößert. Dadurch erfährt das Mitfühlen eine ästhetische Qualität, weil das sinnliche Empfinden auf andere Subjekte geweitet, neue Perspektiven eröffnet⁵². Die Weltwahrnehmung wird durch ein gesteigertes Empfinden von Empathie reicher und gleichzeitig vielschichtiger. Das Mit-Erleben von fremden Perspektiven kann Nähe, Vertrauen und damit eine neue Bindung schaffen. In einem umwelttugendethischen Ansatz ist das Mit-Erleben natürlich nicht auf Mitmenschen beschränkt. Das Mitfühlen mit dem Empfinden der Fürsorge baut schließlich die Brücke zwischen dem Ethischen und Ästhetischen⁵³.

Gleichzeitig ermöglicht dies bereichernde Erleben Mäßigung im eigenen Leben einzuüben, da ein empathisches Zusatzerleben jenes überkompensiert. Dies ist eine Perspektive zu einem umweltentlastenden Verhalten im Individuellen. Ein weiterer Vorzug eines tugendethischen Ansatzes liegt darin, den Kreis der moralisch zu berücksichtigenden Entitäten nicht bestimmen zu müssen (Anthropozentrismus, Pathozentrismus, Biozentrismus, Holismus) und damit das Demarkationsproblem in der Sichtweise der Umweltethik umgehen zu können. Denn eine enge anthropozentrische Position lässt sich darlegen, die nur das eigene Individuum umfasst (Egoismus), erweitert um alle Personen (Personalismus bzw. Humanismus, alle Menschen gegenwärtig und zukünftig), schließlich lässt sich die Sphäre der moralisch zu berücksichtigenden Entitäten um alle leidensfähigen Tiere (Pathozentrismus), alle Lebewesen (Biozentrismus) und um die gesamte Natur (Holismus) erweitern. Der Holismus plädiert dafür, allen Entitäten der Natur einen eigenen Wert zuzusprechen, die Natur ist um ihrer selbst willen zu schützen⁵⁴. Bezogen auf

⁵² Vgl. Breithaupt (2019), S. 206ff.

⁵³ „(Ethik und Ästhetik sind Eins)“ Wittgenstein (1997), 6.421.

⁵⁴ Krebs (1997), S. 342ff., Gorke (2010), S. 23f.

das Abgrenzungsproblem scheint Gorke einzig die umfassendste Position des Holismus selbstevident zu sein, „die Antwort des Holismus ist [...] die] einzige, [die] keiner weiteren Erläuterung bedarf“⁵⁵. Damit glaubt Gorke der Begründungslast, welchen Entitäten ein Eigenwert zugeschrieben werden kann, zu entgehen, denn alle anderen Konzepte der Umweltethik müssen ihrerseits schlüssig darlegen können, warum sie bestimmte Entitäten aus dem Kreis der moralisch zu berücksichtigenden Wesen ausschließen. Man müsste folglich Objekten einen eigenen Wert zusprechen, der nicht aus menschlichem oder animalischem Bewusstsein herleitbar ist und damit einen Gegenbegriff zum instrumentellen Wert, mithin einen objektiven Wert, darstellt. Denn die Kernfrage, der sich anthropozentrische Argumente stellen müssen, ist ob es

wirklich angemessen [ist ...], den mehr als drei Milliarden Jahre alten Prozess der biologischen Evolution und der Selbstorganisation von Ökosystemen vollständig dem Interessenkalkül von *Homo sapiens* unterzuordnen⁵⁶.

Doch auch für den Holismus bleibt die Gefahr bestehen, dass eine anthropozentrische Sicht implizit auf nicht-menschliche Entitäten erweitert wird und letztlich doch eine Zuschreibung von menschlichen Eigenschaften erfolgt. In einer tugendethischen Perspektive ist das Abgrenzungsproblem jedoch nicht zentral, daher lassen sich solche Zuschreibungen wie auch moralische Statuszuschreibungen leicht umgehen. Im Zentrum einer Tugendethik steht eben eine handelnde Person, die aus eudaimonistischen Gründen motiviert ist. Zentral für einen tugendethischen Ansatz ist die individuelle Handlung und gerade im Umweltdiskurs hat sich die Diskrepanz von Wissen und Handeln als zentraler Angelpunkt herausgestellt. Denn in Umweltentscheidungen treten immer wieder dilemmatische Situationen wie den oben geschilderten auf. Jede Handlung zeitigt Schadnabenfolgen und eine handelnde Person verursacht Umweltschäden. Dies spricht für einen tugendethischen Ansatz, welcher aus der Stärkung relevanter Tugenden Handlungswirksamkeit ableitet. Im Umweltverhalten ist

⁵⁵ Gorke (2006), S. 97.

⁵⁶ Ebd., S. 95, ähnlich bei Rolston (1997), S. 264.

ein Maßhalten eine zentrale Tugend, die sich aus regelmäßigen Naturkontakten stärken lässt⁵⁷.

4.3. Die Sinndimension

Menschen sind, anders als KI, sinnbedürftige Wesen und müssen in ihrem Wesen berührt werden, um zu Handeln. Daher bleibt das zunehmende Umweltwissen, der ständige Zustrom ökologischer Daten zu ansteigenden Treibhausgaskonzentrationen und abnehmender Artenvielfalt vielen Menschen äußerlich und betrifft nicht deren existenzielles In-der-Welt-Sein. Morton spitzt das zu: „der Datenmüllmodus verstärkt lediglich das Unvermögen der Dinge, noch etwas für uns zu bedeuten“⁵⁸. Menschen müssen folglich in ihrer Sinndimension als zentrales Motivationsmoment angesprochen werden, denn:

Es gibt wahrscheinlich kein Übel, das der Mensch nicht zu ertragen bereit wäre, wenn er einen Sinn dieses Erleidens einzusehen vermöchte; sicher aber gibt es kein irdisches Gut, dessen Genuß dem Menschen nicht auf die Dauer schal würde, wenn er das Festhalten daran nicht als sinnvoll empfinden könnte⁵⁹.

Sinn lässt sich jedoch nicht über Algorithmen simulieren und stellt daher ein weiteres Distinktionsmerkmal menschlicher und künstlicher Intelligenz dar.

Neben dem Sinn, den ein Individuum für sich entdecken kann, gilt es „Naturwesen außerhalb des Menschen“, deren „Seins- und Sinnentwürfe, [...] die uns sittlich verpflichten“ zu berücksichtigen.⁶⁰ Denn das menschliche Dasein ist wesentlich durch Beziehungen, damit relational konstituiert. Daher stehen Menschen dem nicht-menschlichen Leben in „einer Solidarität der Sinnerwartung mit allem Lebendigen; eine Solidarität, die in symphonischer Resonanz von uns gespürt wird“⁶¹ gegenüber. „Der Mensch kann nur ganz Mensch werden, wenn er alle Dinge nicht nur utilitär

⁵⁷ Vgl. Verf. (2020a, b, c).

⁵⁸ Morton (2019), S. 238.

⁵⁹ Hengstenberg (1980), S. 7.

⁶⁰ Ebd., S. 48.

⁶¹ Ebd., S. 49.

„nimmt“, sondern sie auch in ihrem Eigensein konspirierend „ver-nimmt“ um ihrer selbst willen“⁶². Für Hengstenberg entspricht dies dem Gebot der Sachlichkeit, sich nichtmenschlichen Naturwesen um ihrer selbst willen zuzuwenden. Damit entsteht für ihn eine „Verbindlichkeit im Verhältnis zu allem Lebendigseiden, nicht nur zum Mitmenschen“.⁶³ Dieses „universelle Sinngebot“ zeigt sich als eine tragfähige Basis, Natur als Quelle für menschlichen Sinn auszubauen. Dabei erschließt sich die Tiefe des Sinns nicht in einem fortwährenden „Mehr“, die Sinnerschließung bedarf des Mutes innezuhalten und zu erkennen, dass die Konzentration darauf, weniger zu sehen, weniger zu erleben, weniger zu unternehmen, die Qualitäten des Wenigen steigert und das Bedürfnis nach immer mehr senkt. Ein tugendethischer Ansatz ist an eine vermeintliche Einschränkung des Individuums gebunden, die sich jedoch als qualitativer Zugewinn herausstellt. Denn im Vielen liegt Schnelligkeit, Oberflächlichkeit, Beliebigkeit, Gewinn jedoch lässt sich aus einer qualitativen Beziehung schöpfen. Merton illustriert den Gedanken durch einen Museumsbesuch:

Ein Reisender kann mit dem Baedeker durch ein Museum gehen, gewissenhaft alles Wichtige betrachten und doch unbefriedigter herauskommen, als er hineingegangen ist. Alles hat er betrachtet und nichts gesehen. Er hat sich unendlich angestrengt und hat sich ermüdet. Hätte er innegehalten, um ein einziges Bild zu betrachten, das ihn wirklich ansprach, und darüber alles andere beiseitegelassen, so könnte er sich mit dem Gedanken trösten, seine Zeit nicht gänzlich verschwendet zu haben. Er würde etwas in sich selbst entdeckt haben, nicht bloß außer sich⁶⁴.

KI könnte die Informationsfülle des Museums strukturieren, jedoch ohne Aussicht auf Sinn oder einen Beitrag zu einem guten Leben. Es bleibt der natürlichen Intelligenz vorbehalten, die qualitative Dimension im Leben zu stärken und daraus Handlungsmotivation abzuleiten, denn: „Die meisten von uns wissen oder ahnen im Innersten recht genau, was wert- und sinnvoll wäre und was nicht. Woran es aber bislang zu mangeln scheint, ist das Wissen, wie man konkret und realistisch wert- und sinnorientiert engagiert leben kann; und auch das Wissen

⁶² Ebd., S. 50.

⁶³ Ebd., S. 50.

⁶⁴ Merton (1984), S. 121f.

darum, dass sinnorientiertes, verantwortliches Handeln nicht nur die Welt bereichert, sondern auch uns selbst“.⁶⁵ Auch dies setzt einer konsequentialistischen Denkweise Grenzen, die sich überwinden lassen durch das nachhaltige Handeln, welches in einer Mäßigung des Konsums und Verhaltens von Menschen liegen. Dies schlägt die Brücke von der Sinndimension zu einem umwelttugendethischen Ansatz, die im Umweltdiskurs bislang zu undeutlich aufgetreten ist. Denn eine Grundkonzeption tugendethischer Ansätze liegt darin, dass sich eine Person sich selbst zu als bedeutsam erkannten Tugenden bzw. Handlungen hin entwickelt. Einsichten in den Sinn dieser Tugenden werden dann individuell handlungsleitend da sie eine unmittelbare persönliche Betroffenheit evozieren. Dies kann das, den Umweltdiskurs seit Jahrzehnten begleitenden Lamento über wahlweise das System, den Kapitalismus oder die Großkonzerne durch die Einsicht in die individuelle Handlungsmacht positiv ergänzen. Dadurch entgeht der Einzelne einer Opferrolle und gewinnt Eigenverantwortung und daraus eine weitere Quelle für das gute Leben.

5. Ausblick

KI kann helfen, auf politischer Ebene konsequentialistische Ansätze der Umweltbewertung vorzulegen. Daraus lassen sich Handlungsfolgen bestimmen und entsprechende Grenzwerte in der Ressourcennutzung festlegen. Die Grenzen der KI in der Überwindung der ökologischen Krise liegen darin, dass sie im rein Quantitativen verwurzelt bleibt. Die qualitativen Dimensionen menschlichen Lebens lassen sich jedoch nicht simulieren. Dies gilt auch für den Beitrag der Naturwissenschaften, denn die Naturauffassung, welche empirische Naturwissenschaft und ästhetischen Naturgenuss vereinigte, war bei Alexander von Humboldt noch präsent, ist jedoch in der jüngeren Naturwissenschaft verlorengegangen und diese Entzweiung erscheint sich durch KI noch zu verstärken. Ein umwelttugendethischer Ansatz, der darauf baut, die neuzeitlichen Entzweiungstendenzen von Mensch und Natur zu überwinden, kann in der regelmäßigen Naturbegegnung eine neue Wertschätzung für die Mitwelt, und daraus abgeleitet, ein erhöhtes Engagement für diese erreichen.

⁶⁵ Batthyány (2017), S. 26.

Literatur

- Abram, David (2015): *Im Bann der sinnlichen Natur. Die Kunst der Wahrnehmung und die die Mehr-als-menschliche Welt*, Klein Jasedow
- Al-Chalabi, Malek (2015): Vertical farming: Skyscraper sustainability?, in: *Sustainable Cities and Society* 18, S. 74–77
- Batthyány, Alexander (2017): *Die Überwindung der Gleichgültigkeit. Sinnfindung in einer Zeit des Wandels*, München
- Baumgartner, Christoph (2004): *Umweltethik – Umwelthandeln. Ein Beitrag zur Lösung des Motivationsproblems*, Paderborn
- Belton, Valerie/Stewart, Theo (2002): *Multiple Criteria Decision Analysis. An Integrated Approach*, Dordrecht
- Bendel, Oliver, Hg. (2018): *Pflegeroboter*, Wiesbaden
- Berry, Wendell (2002): *Körper und Erde. Essay über gutes Menschsein*, Klein Jasedow
- Bloom, Paul (2018): *Against Empathy. The Case for Rational Compassion*, London
- Breithaupt, Fritz (2019): *Die dunklen Seiten der Empathie*, Berlin
- Devall, Bill (1997): Die tiefenökologische Bewegung, in: *Ökophilosophie*. Hg. von Dieter Birnbacher, Stuttgart, S. 17–59
- Dreitzel, Hans P. (1992): *Reflexive Sinnlichkeit. Mensch Umwelt Gestalttherapie*, Köln
- Goleman, Daniel (2009): *Ökologische Intelligenz. Wer umdenkt, lebt besser*, München
- Hartmann, Nicolai (*1962): *Ethik*, Berlin
- Klages, Ludwig (1914): Mensch und Erde, in: ders., *Sämtliche Werke*, Band 3, Philosophie III, Hg. Ernst Frauchiger u. a., Bonn 1974
- Diakoulaki, Danae u. a. (2005): MCDA and energy planning, in: José Figueira u. a., Hg., *Multiple Criteria Decision Analysis: State of the Art Survey*, Berlin, S. 859–897
- Finger, Robert u. a. (2019): Precision farming at the nexus of agricultural production and the environment, in: *Annual Review of Resource Economics* 11, S. 313–335
- Festinger, Leon (1962): Cognitive dissonance, in: *Scientific American* 207/4, S. 93–107
- Fuchs, Thomas (2020): *Verteidigung des Menschen. Grundfragen einer verkörperten Anthropologie*, Berlin
- Gehlen, Arnold (*1978): *Der Mensch. Seine Natur und seine Stellung in der Welt*, Wiesbaden
- Gorke, Michael (2010): *Eigenwert der Natur*, Stuttgart
- Halla, Michael u. a. (2018): Believing in climate change, but not behaving sustainably: Evidence from a one-year longitudinal study, in: *Journal of Environmental Psychology* 56, S. 55–62
- Hengstenberg, Hans-Eduard (1980): *Sinn und Sollen. Zur Überwindung der Sinnkrise*, Essen

- Konrath, Sara H. u. a. (2011): Changes in dispositional empathy in American college students over time: A meta-analysis, in: *Personality and Social Psychology Review* 15, S. 180–198
- Krebs, Angelika, Hg. (1997): *Naturethik*, Frankfurt am Main
- Kuhn, Helmut (1981): *Der Weg vom Bewußtsein zum Sein*, Stuttgart
- Lipps, Theodor (²1907): *Fühlen, Wollen und Denken. Versuch einer Theorie des Willens*, Leipzig
- Lübbe, Hermann (2019): *Politischer Moralismus. Der Triumph der Gesinnung über die Urteilskraft*, Münster
- Merton, Thomas (²1984): *Keiner ist eine Insel*, Zürich
- Mustajoki, Jyri/Marttunen, Mika (2017): Comparison of multi-criteria decision analytical software for supporting environmental planning processes, in: *Environmental Modelling & Software* 93, S. 78–91
- Nietzsche, Friedrich (1999): Also sprach Zarathustra, in: ders., *Kritische Studienausgabe*. Hg. von Giorgio Colli & Mazzino Montinari, Bd. 4, München
- Rathmann, Joachim (2020a): Gesundheitsressource Landschaft, in: Jens Soentgen u. a., Hg., *Umwelt und Gesundheit*, Baden-Baden, S. 167–197
- (2020b): Von der Naturkunde zur Umwelttugendethik: Ein möglicher Weg zur Überwindung der Diskrepanz von Umweltwissen und Umwelthandeln?, in: *Comenius-Jahrbuch* 28, S. 97–120
- (2020c): *Therapeutische Landschaften. Landschaft und Gesundheit in interdisziplinärer Perspektive*, Wiesbaden
- Rolston, Holmes III (1997): Werte in der Natur und die Natur der Werte, in: *Naturethik*. Hg. von Angelika Krebs, Frankfurt am Main, S. 247–270
- Schrems, Isabel/Upham, Paul (2020): Cognitive Dissonance in Sustainability Scientists Regarding Air Travel for Academic Purposes: A Qualitative Study, in: *Sustainability* 12, S. 1837; doi:10.3390/su12051837
- Schrott, Raoul (2011): *Homer Ilias, übertragen von Raoul Schrott*. Frankfurt am Main
- Spranger, Eduard (1962): *Gedanken zur Daseinsgestaltung. Ausgewählt von Hans Walter Bähr*, München
- Stammel, Barbara/Cyffka, Bernd (2015): *Naturschutz*, Darmstadt
- Steffen, Will u. a. (2015): The trajectory of the Anthropocene: The Great Acceleration, in: *The Anthropocene Review* 2/1, S. 81–98
- Szymborska, Wislawa (2014): *Glückliche Liebe und andere Gedichte*, Berlin
- Twenge, Jean M. u. a. (2018): Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology, in: *Emotion* 18/6, S. 765–780
- Vetter, August (1949): *Natur und Person. Umriss einer Anthropognomik*, Stuttgart

Waal, Frans de (2011): *Das Prinzip Empathie. Was wir von der Natur für eine bessere Gesellschaft lernen können*, Darmstadt

Wittgenstein, Ludwig (¹¹1997): *Tractatus logico-philosophicus*, in: ders., *Werkausgabe* Bd. 1, Frankfurt am Main

„Wer nichts als Chemie versteht, versteht auch die nicht recht.“

Interdisziplinarität als Mehrsprachigkeit¹

Jens Soentgen

Zusammenfassung

Interdisziplinarität ist wissenschaftliche Mehrsprachigkeit, nämlich die Fähigkeit, sich im Wissenschaftssystem über die Grenzen der eigenen Disziplin hinaus verständigen zu können. Diese Art der Mehrsprachigkeit zu fördern ist heute aus zwei Gründen dringend erforderlich: zum einen wegen der wachsenden Spezialisierung und zugleich gegenseitigen Verflechtung der einzelnen Disziplinen und zum anderen wegen der Situation im Anthropozän, in dem sich Phänomene nicht mehr nur aus der Perspektive einer einzigen Disziplin erklären lassen und insbesondere auch geistes- und kulturwissenschaftliche Aspekte zu berücksichtigen sind.

Abstract

Interdisciplinarity is academic multilingualism, namely the ability to communicate across the boundaries of one's own discipline within the system of the sciences. To foster this kind of multilingualism, today is urgently required because of two reasons: on the one hand because of the growing specialization and at the same

¹ Erweitertes Manuskript eines Vortrags, den ich zuerst im Oktober 2019 auf Einladung der interdisziplinären Studienschwerpunkte (iSp) der TU Darmstadt in der Schader-Stiftung hielt. Teile dieses Textes wurden veröffentlicht in Verf. (2021).

time increasing interconnection between the single disciplines, and on the other hand because of the situation in the Anthropocene, in which phenomena cannot be explained any more from the perspective of only one discipline alone and especially aspects of the humanities have to be taken into consideration.

1. Einleitung

In eines seiner Sudelbücher² schreibt Lichtenberg folgende Notiz:

Rousseau hat glaube ich gesagt: ein Kind, das bloß seine Eltern kennt, kennt auch die nicht recht. Dieser Gedanke läßt sich auf viele andere Kenntnisse, ja auf alle anwenden, die nicht ganz reiner Natur sind: Wer nichts als Chemie versteht, versteht auch die nicht recht.

Die Aufzeichnung stammt vermutlich aus dem Jahr 1790. Man ist allzu leicht geneigt, Lichtenberg einfach zuzustimmen, natürlich, wer nichts als Chemie versteht, versteht auch die nicht recht. Weil er sie, so sagen wir uns, nicht im Kontext versteht, wie auch immer der beschaffen sein mag. Doch Lichtenbergs Aphorismus hat eine gewisse Tiefenstruktur. Lichtenberg macht nämlich eine Ausnahme: Kenntnisse, die ganz rein sind. Die kann man offenbar tatsächlich ganz verstehen, auch wenn man nur sie versteht. Zwar wissen wir nicht, woran Lichtenberg dachte, ich nehme aber an, an die Mathematik. Die beruht auf Axiomen, die gesetzt werden, und sollte daher rein aus sich selbst heraus verständlich sein.

Die Chemie aber, und mit ihr fast alle anderen Wissenschaften, ist nicht so rein. Sie benötigt das Experiment, sie benötigt Stoffe, die sie irgendwoher nehmen muss, die ihr angeliefert werden. Diese Stoffe macht sie teils selbst, teils kommen aus Bergwerken, fernen Gegenden, aus Pflanzen und anderen Lebewesen. So ergibt sich sofort, dass andere Kenntnissysteme in sie eindringen, die Geologie, die Medizin, die Botanik und allgemein die Biologie, wie wir heute sagen würden. Heute scheint uns die Chemie ein monolithisches, rein selbstreferentielles System zu sein, mit den imposanten Symbolen des Periodensystems und der Benzolformel, die August Kekulé erfand, und dessen Geheimnisse nur Eingeweihten zugänglich sind. Doch

² Heft J 838, zitiert nach: Lichtenberg (1983), S. 408.

Lichtenbergs Aphorismus behauptet etwas anderes: Sie ist ein Polysystem, in dem sich verschiedene Wissenssysteme durchdringen. Sie ist, mit anderen Worten, in sich selbst schon interdisziplinär. Dieses Fach ist viele Fächer. Deshalb versteht sie nur, wer mehr versteht als sie.

Dies ist auch den meisten Chemikern völlig klar. Ein Chemiker würde nie, und darin unterscheidet sich die Chemie positiv von der Durkheim-Schule in der Soziologie, ein Chemiker würde also nie die Regel aufstellen, dass chemische Fakten nur durch andere chemische Fakten erklärt werden können. Hieße das doch, dass er die Tatsachen des Periodensystems, die nur durch die Quantenmechanik erklärt werden können, ignorieren müsste. Auch das Vorkommen von bestimmten Eiweißen und Fetten in Organismen könnte ein Chemiker dann nicht erklären, sind dies doch biologische Tatsachen. Was die DNA leistet, und warum sie so wichtig ist, ist ebenfalls nur durch nicht-chemische Fakten erklärbar.

Chemiker wissen das auch ganz genau, ja mehr noch, sie sind sehr pragmatisch und erfinderisch darin, aus nicht-Chemie Chemie zu machen, etwa aus Quantenmechanik eine Theorie der chemischen Bindung, die zwar für Physiker viel zu oberflächlich ist, in der chemischen Forschung aber hervorragende Dienste leistet. Insofern funktioniert die Interdisziplinarität mancherorts ganz gut.

Anderswo ist die Lage allerdings nicht ganz so gut, wenn man etwa nach Kontakten zwischen Chemie und den Kultur- und Gesellschaftswissenschaften fragt. Zwischen diesen Disziplinen finden wir nicht gerade, wie zwischen Chemie und Physik oder Chemie und Informatik oder Chemie und Biologie gewaltige mehrspurige Autobahnen. Da sind es eher kleine, kaum mehr erkennbare, oft schon überwachsene Trampelpfade und ungepflegte, verfallene Wege, die nur von einzelnen Waldläufern und Individualisten, manche würden sagen, von Spinnern begangen werden. Nun, unter diesen Spinnern sind zwar durchaus bedeutende Leute, war doch der Biochemiker Joseph Needham (1900–1995) zugleich ein sehr bedeutender Sinologe. Ein anderer Chemiker, genauer Chemieingenieur, Benjamin Lee Whorff (1897–1941), war ein bedeutender Sprachphilosoph.

Insgesamt aber tun sich Chemiker, wie auch andere Naturwissenschaftler schwer mit den Gesellschafts- und den Geisteswissenschaften. Ein *obstacle épistémologique*, um Gaston Bachelard³, den französischen Philosophen und Chemiker zu zitieren, ein Erkenntnishindernis, das die Interdisziplinarität über den *great divide*, über den

³ Bachelard (1972).

großen Graben erschwert, ist der Chauvinismus, die Vorstellung, dass nur die eigene Wissenschaft eine wirkliche Wissenschaft sei.

Hans Georg Gadamer hat mir dies bei einem Gespräch, das ich vor vielen Jahren mit dem damals fast hundertjährigen in Heidelberg führen durfte, einmal recht drastisch erklärt. Sein Vater war Chemiker, und als ihm Hans Georg eröffnete, er wolle Philosophie studieren, sagte dieser: „Dass Du mir das antust, zu den Schwatz-Professoren zu gehen!“ Ich glaube, auch heute noch denken nicht wenige Chemiker ähnlich. Was die anderen sprechen, ist keine eigentliche Sprache, sondern nur vornehmes Ornament, das in erster Linie soziale Funktionen hat, es fehlt der Sachbezug, so glaubte Gadamer Senior damals. Im Grunde eine Variante des Barbaren-Konzeptes der alten Griechen, die die Nichtgriechen als Barbaren bezeichneten, weil diese keine eigentliche Sprache sprachen, sondern nur Bar-Bar, oder bla-bla, wie wir heute sagen würden. Diese Kommunikationsverweigerung zwischen gerade der Chemie und den Geisteswissenschaftlern hat aber keine systematischen Gründe, denn alle modernen Wissenschaften haben eine gemeinsame Geschichte, gehen von ähnlichen Methoden aus, streben nach empirisch überprüfbaren Resultaten. Sie sind und müssen untereinander kooperationsfähig sein. Wo Kooperation nicht möglich zu sein scheint, da liegen keine prinzipiellen Gründe vor, die hier bestehende Sprach- und Weglosigkeit ist vielmehr historisch gewachsen. Meiner Meinung nach handelt es sich um eine Altlast des 19. Jahrhunderts, als für eine recht kurze Zeit Philosophen der Meinung waren, der Chemie die Methode vorschreiben zu können. Wir müssen diese Altlast heute abtragen, wieder Wege freiräumen zwischen Chemie und Gesellschaftswissenschaften, zwischen Chemie und Geisteswissenschaften.

Erst wenn das geschieht, können wir auch die Chemie selbst verstehen, denn genau diese Naturwissenschaft hat – gerade in Deutschland – eine politische Rolle, sie war für das deutsche ‚Nationbuilding‘, wie sich an vielen Beispielen zeigen ließe, vom Rübenzucker bis zum Kunstgummi, unerlässlich. Deutschland hatte bekanntlich als zu spät gekommene Nation zunächst und nach 1918 dann wieder keine Kolonien, aber genau hier bot die herausragende deutsche Chemie einen Weg: Was andere als Indigo oder Kautschuk oder Kampfer in ihren tropischen Kolonien von den dort unterworfenen Kolonialvölkern abpressten, das erzeugte das damalige Deutsche Reich aus der Luft, aus Kohle und Kalk oder zur Not auch aus der Bierhefe. Eine zunächst sehr erfolgreiche Modernisierungsstrategie⁴. Es wäre, so zeigt

⁴ Verf. (2014).

diese Verflechtung von Politik und Chemie, auch für Chemikerinnen und Chemiker sinnvoll, sich von vorn herein in einem umfassenderen Sinn mit Geschichte und Politik zu beschäftigen, auch um sich selbst besser zu verstehen. Zudem könnte ein solches breiteres Studium dazu beitragen, nicht so rasch auf jene Abwege zu geraten, wie etwa jene, die sich in der NS-Zeit gleichgeschaltet haben.

2. Interdisziplinarität

Mit dem Beispiel sind wir am systematischen Ort von Interdisziplinarität angekommen. Sie soll historisch gewachsene Hindernisse oder die Sprachlosigkeit zwischen bestimmten Disziplinen beseitigen. Interdisziplinarität reagiert auf ein Problem.

Aber was genau verstehen wir unter Interdisziplinarität? Zunächst einmal handelt es sich um ein sehr sperriges Wort, zu dem sich inzwischen eine umfangreiche Literatur gebildet hat. Man unterscheidet Multidisziplinarität, Transdisziplinarität, Pluridisziplinarität, Crossdisziplinarität⁵; bei Niklas Luhmann findet man okkasionelle, temporäre und transdisziplinäre Interdisziplinarität⁶ und diese Worte sind nicht nur Zungenbrecher, sie geben auch Rätsel auf, was mit ihnen eigentlich gemeint sei.

Das Präfix *inter* klingt immer gut, wenn auch etwas angestaubt, wie aus der Zeit von Interlücke und Interatom, es klingt nach den frühen Siebzigern. Manche warnen ohnehin davor, dass vieles, was interdisziplinär sei, eigentlich subdisziplinär wäre, wenn nicht a-disziplinär. Auch Pseudo-Interdisziplinarität, zusammengesetzte, unterschiedslose oder Hilfs-interdisziplinarität werden kritisiert. Es sind inzwischen so viele Konzepte und Modelle, dass Uwe Voigt den Diskurs mit einem „Modell der Modelle“⁷ geordnet hat. Als kürzlich in der Zeitschrift *Gaia* ein weiterer konzeptioneller und sehr differenzierter Artikel zu der Frage Was ist Transdisziplinarität erschien⁸, meldete sich in einer Replik der Philosoph Jürgen Mittelstraß, der Schöpfer des Transdisziplinaritätsprinzips. Er geißelte den wuchernden Transdisziplinaritätsdiskurs und meinte, der Diskurs verdunkle inzwischen, statt zu erhellen⁹.

⁵ Jungert u. a. (2013).

⁶ Luhmann (1990) S. 46–462.

⁷ Voigt (2013).

⁸ Krohn/Grunwald/Ukowitz (2017).

⁹ Mittelstraß (2018). Vgl. auch ders. (1994), S. 108, zu Transdisziplinarität.

Die Diskussion ist inzwischen so verwirrend geworden, dass ich einen Neuanfang vorschlagen möchte. Dabei soll kein neues begriffliches Konstrukt für Interdisziplinarität eingeführt werden. Es ist ja nicht so, dass das begriffliche Denken das einzig produktive wäre. Auch das Erfinden von Bildern¹⁰ kann uns weiterbringen oder das Erzählen von Geschichten. Und auch einfache Analogien sind nicht zu verachten. Da bin ich ganz einig mit Lichtenberg, der den Vergleich zu seiner Haupt-Denktechnik erklärte. Für Lichtenberg ist der witzige Vergleich bekanntlich, wie er in einem weiteren Vergleich sagt, das Gegenteil des Mikroskopierens, sei kein Vergrößern, sondern ein Verkleinern, und er behauptet, dass nicht alle Entdeckungen mit Vergrößerungsgläsern gemacht werden können, sondern auch viele mit Verkleinerungsgläsern¹¹. Dieses Verkleinern, das der Vergleich ermöglicht, hat den Vorzug, dass Zusammenhänge sichtbar werden. Daher sind Analogien, die kognitiv wichtige Zusammenhänge herstellen, bei unserem Thema von Bedeutung. Üben wir uns also in der Vergleichskunst, suchen wir Ähnlichkeiten! Dies ist, wie viele Lichtenberg-Forscher hervorgehoben haben, der Kern des Lichtenbergschen Metaphorisierens¹².

3. Wer nur seine eigene Sprache versteht, versteht auch die nicht recht

Ich möchte also eine Analogie vorschlagen, auch wenn ich natürlich kein Lichtenberg bin. Die Analogie ist daher wenig originell, sie wird aber vielleicht helfen, das, worum es uns in der Interdisziplinarität geht, klarer zu machen und auch neue Wege dorthin weisen. Zunächst möchte ich die Disziplinen, die Fächer, die wir heute haben, mit Sprachen vergleichen. Das ist schon öfter so geschehen. Es passt auch quantitativ, denn bereits Ende der 1980er Jahre zählte man über 4000 verschiedene Fächer. Inzwischen dürften noch einige hinzugekommen sein¹³. Und wir haben auch ungefähr 7000 Sprachen.

Das Entstehen ganz spezifischer Fachsprachen in der Wissenschaft ergibt sich aus der Arbeitsteilung in der Wissenschaft, die natürlich mit der Ausbildung von sozia-

¹⁰ Siehe zur kognitiven Leistungsfähigkeit von Bildern Köller (2012), besonders S. 106–128.

¹¹ Sudelbücher, Nr. D. 469, in: Lichtenberg (1983), S. 161.

¹² Neumann (1976), S. 141f.

¹³ Mittelstraß (1989), S. 68.

len Gruppen einhergeht. Chemiker, Biologen, und weiter: Organiker, Anorganiker usw. Und ähnlich in der Philosophie, wo es die analytischen Philosophen gibt, die Ethiker, die Phänomenologen und so weiter. Die Arbeitsteilung, die in der Wissenschaft wirksam ist, die erreicht längst die einzelnen Fächer. Es bilden sich Gruppen. Und diese Gruppen verhalten sich nicht anders als Gruppen überall. Sie setzen sich nämlich, wie Ludwik Fleck und viele andere gezeigt haben, mit ihren eigenen Denkweisen und Sprachen kontrastiv von anderen Gruppen ab, um die eigene Gruppenidentität zu festigen. Wir sind nicht die, sondern die. Wir reden nicht nur, wir experimentieren auch. Oder: Wir rechnen, ihr diskutiert. Es gab einmal einen baden-württembergischen Ministerpräsidenten, der die Geistes- und Gesellschaftswissenschaften insgesamt als „Diskussionswissenschaften“ von den „Produktionswissenschaften“ zu unterscheiden für gut befand.

Hier wirkt sich die Vorstellung aus: Was die Gruppenmitglieder tun, ist anders und meist nach eigener Einschätzung besser als das, was alle anderen tun. Und auch die Sprachen differieren, nicht nur aufgrund sachlicher Notwendigkeit. Die Menschen wollen nicht nur etwas sagen, sie wollen es auch anders sagen. Seit alter Zeit beobachten die Menschen die Entstehung von Sondersprachen und die Probleme, die daraus erwachsen. Daher gibt es auch zahlreiche Geschichten, die wir uns mit unserer Metapher erschließen können. Unter diesen Geschichten ist wohl die bekannteste die vom Turmbau zu Babel aus dem Alten Testament (Genesis 11, 1–9). Worum geht es dort: Die Bibel erzählt von einem Volk aus dem Osten, das „die eine“ (heilige) Sprache spricht und sich in der Ebene in einem Land namens Schinar ansiedelt. Dort will es eine Stadt und einen Turm „mit einer Spitze bis zum Himmel bauen“. Da stieg der Herr herab, um sich Stadt und Turm anzusehen, die die Menschenkinder bauten. „Nun befürchtet er, dass ihnen nichts mehr unerreichbar sein [wird], was sie sich auch vornehmen“, dass das Volk übermütig werden könnte und vor nichts zurückschreckt, was ihm in den Sinn kommt. Gott „verwirrt“ ihre Sprache und vertreibt sie „über die ganze Erde“. Die Weiterarbeit am Turm endet gezwungenermaßen, weil die durch Eingreifen Gottes aufgetretene Sprachverwirrung die notwendige Verständigung der am Turm bauenden Menschen untereinander so gut wie unmöglich macht.

Die Bezüge zu unserem Thema sind offensichtlich¹⁴. In der Geschichte vom Turmbau zu Babel geht es um Arbeitsteilung, durch die gewaltige Werke möglich

¹⁴ Siehe zur Interpretation der Geschichte vom Turmbau zu Babel Köller (2006), S. 91–120.

werden. In dieser Geschichte ist die Gefahr, die mit der Spezialisierung einhergeht, benannt, denn das Resultat der hier ganz plötzlich, nicht allmählich auftauchenden Sprachenvielfalt ist, dass das gemeinsame Werk verlassen wird und die Menschen sich über die Erde verstreuen. Hier besteht ein kausaler Zusammenhang: Wo jeder nur noch seine eigene Fachsprache, nur sein eigenes Gewerk versteht und hören kann, wo die erworbene Schwerhörigkeit gegenüber den anderen immer weiter zunimmt, da zerfallen Gesellschaften. Oder umgekehrt: Wo es Arbeitsteilung gibt, muss es auch Verständigung geben. Wer dazu nicht bereit ist, der kann nicht kooperieren, übrigens auch in seinem eigenen Fach nicht. Wenn also jemand sagt, Kooperation sei schwierig und zeitaufwendig, sollte man antworten: Kooperation ist überhaupt schwierig und zeitaufwendig, ob interdisziplinär oder nicht. Aber ohne Kooperation ist Wissenschaft, jedenfalls moderne Wissenschaft unmöglich. Die Probleme mit der Interdisziplinarität sind nur gesteigerte Probleme dessen, womit wir uns auch sonst auseinandersetzen. Und damit kommen wir zu folgender Schlussfolgerung: Die zentrifugalen Tendenzen, die mit Arbeitsteilung, ob auf dem Bau oder in der Wissenschaft, einhergehen, müssen immer wieder aufgehalten werden, weil sonst auch die Vorteile der Arbeitsteilung verloren gehen. Es muss immer eine gemeinsame Verständigungsebene geben. Das gilt beim Turmbau ebenso wie in der Wissenschaft. Das bedeutet: Interdisziplinarität muss konsequent kultiviert werden, sie ist nicht mit einem Male da und bleibt. Man sagt ja auch nicht: Die deutsch-französische Freundschaft ist nun ein für alle Male etabliert. Nein, sie muss immer wieder neu gepflegt werden. Interdisziplinarität ist überhaupt kein Zustand, sondern ein dauernder Prozess, der sich nicht abschließen lässt.

4. Interdisziplinarität als Mehrsprachigkeit: Was bedeutet das?

„Wer fremde Sprachen nicht kennt, weiß nichts von seiner Eigenen“. Diese Variante des Lichtenberg-Spruches stammt von Goethe, sie findet sich in seinen 1833 erschienen *Maximen und Reflexionen*¹⁵ und ich vermute, dass er sich ebenso wie Lichtenberg ursprünglich von Rousseau hat anregen lassen.

¹⁵ Zitiert nach Wandruszka (1979), S. 333.

Die Vorteile der Mehrsprachigkeit kennt man. Auch wenn man nur rudimentäre Kenntnisse einer Fremdsprache hat, erschließt man sich doch eine völlig neue Welt. Man kann radebrechen, erfindet sich eine eigene Behelfssprache, die aber ganz neue Handlungen ermöglicht. Und mehr noch: Man versteht halbwegs, was die anderen schreiben, vielleicht auch, was sie sprechen. Wenn jeder Wissenschaftler das B2-Niveau in einer von seinem Fach aus möglichst weit entfernten anderen Disziplin erreichen würde, hätten wir die Probleme mit mangelnder Interdisziplinarität nicht. Die Universitäten wären dann nicht mehr Orte monodisziplinärer Uniformität, sondern tatsächlich mehrsprachiger Universalität.

Interdisziplinarität ist Mehrsprachigkeit, das ist also die Analogie, auf die ich hinauswill. Ein bisschen mehrsprachig sind wir alle, schon alleine deshalb, weil die weitaus meisten Kenntnissysteme, um mit Lichtenberg zu sprechen, keine Monosysteme sind, sondern Polysysteme. Jedes Fach ist viele Fächer. Ein wenig Mehrsprachigkeit gibt es sogar über die großen Fächerkulturen hinweg, über den Block der Naturwissenschaften auf der einen und der Gesellschafts- und Geisteswissenschaften auf der anderen Seite hinweg. Auch Naturwissenschaftler haben eine grobe Vorstellung, wie Historiker z. B. arbeiten oder Kunstwissenschaftler. Zumindest kennen sie die Chronologie ihres eigenen Faches.

Wir können ein Grundverständnis aufbauen, das für wissenschaftliche Kooperationen ausreicht, und dazu, die Beiträge der anderen Disziplinen, ihre Potentiale zu würdigen und einzubeziehen. Natürlich: Unsere Mehrsprachigkeit wie auch unsere Interdisziplinarität wird immer unvollkommen sein. Aber auch die eigene Muttersprache spricht niemand perfekt. Und das sollte uns ermutigen! Wissenschaftler unter anderen Wissenschaftlern zu sein bedeutet in einer immer unvollkommenen Mehrsprachigkeit zu leben.

Eine gewisse Mehrsprachigkeit gibt es also auch in den spezialisiertesten Disziplinen. Anders könnte moderne Forschung auch nicht funktionieren. Doch diese Mehrsprachigkeit ist eben auch der Steigerung fähig, und eine solche Steigerung ist ausgesprochen wünschenswert, wenn wir die großen wissenschaftlichen Aufgaben unserer Zeit angehen wollen. Dabei geht es nicht darum, dass wir alle gewissermaßen Sprachengenies werden sollen. Aber es ist auch zu wenig, wenn jeder sich nur mit den Fächern beschäftigt, die dem eigenen Fach am nächsten sind, der Chemiker z. B. mit Physik oder der Soziologe noch ein wenig mit Psychologie.

Nun werden manche einwenden, dass sie ja dann nur Dilettanten sind, wenn sie versuchen, sich mit anderen als den eigenen Fächern zu befassen. Doch auch in un-

seren eigenen Fächern sind wir bis zu einem gewissen Grad Amateure. Und genauso in unserer Muttersprache, denn niemand kann von sich behaupten, dass er oder sie so gut Deutsch kann, dass seine oder ihre muttersprachliche Kompetenz nicht mehr verbessert werden könne. Das Dazulernen bleibt eine lebenslange Aufgabe.

Ein anderer oft zu hörender Einwand lautet: ‚Ich habe kaum Zeit, mich mit dem eigenen Fach so zu befassen, wie ich es müsste, wie soll da noch Zeit für andere Fächer übrigbleiben.‘ Auch hier verschafft uns die Analogie eine neue Perspektive. Denn man kann, wie der Sprachforscher Wandruszka herausgearbeitet hat, unterscheiden zwischen der ausübenden und der verstehenden Mehrsprachigkeit. Wandruszka hat auch unterstrichen, dass die verstehende Mehrsprachigkeit viel leichter zu erlangen ist als die ausübende. Schon in unserer Muttersprache verstehen wir viel mehr Wörter, als wir aktiv verwenden. Wir verstehen grundsätzlich, auch in unserer Muttersprache, weit mehr als wir aktiv verwenden, dieses Verhältnis ist, wie Wandruszka schätzt, wie 1000 zu eins¹⁶. Und ein verstehendes Verständnis einer fremden Sprache ist, wie jeder von uns weiß, viel leichter und schneller zu erwerben als ein flüssiger aktiver Gebrauch. Und ganz genauso ist es auch mit dem Verstehen anderer Fächer. Man kann also unterscheiden zwischen der ausübenden und der verstehenden Interdisziplinarität. Und weil das möglich ist, weil es viel rascher geht, zu verstehen, wie andere Fächer arbeiten als sich die Kompetenzen anzueignen, die für diese Arbeit dann tatsächlich erforderlich sind, glaube ich auch, dass die Interdisziplinarität die Wissenschaftler nicht überfordert. Wir alle könnten interdisziplinärer sein als wir sind. Denn es ist leichter und schneller möglich, z. B. chemische oder epidemiologische Forschung in ihren Grundzügen zu verstehen als solche Forschung tatsächlich zu betreiben. Ebenso ist es viel leichter, die Forschungsprinzipien von Historikern (oder von Soziologen) in ihren Grundzügen zu verstehen als solche Forschung dann tatsächlich durchzuführen. Wenn wir aber auch nur *etwas* besser verstehen, was in der Nachbarschaft geschieht, ist schon sehr viel gewonnen, nicht zuletzt erschließen wir uns auch selbst kognitive Ressourcen und werden inspiriert.

Aus der Analogie ergibt sich, wir sahen es bereits, dass Interdisziplinarität wie Mehrsprachigkeit gepflegt werden muss. Sie ist nicht ein für alle Mal angekommen und dann einfach da. Und keine Sorge: Seine Muttersprache wird niemand verlieren. Der Effekt ist vielmehr: Man kann die eigene Muttersprache durch eine Vielzahl neuer Perspektiven bereichern.

¹⁶ Ebd., S. 21.

5. Wer nur seine eigene Wissenschaft versteht, ist kein Wissenschaftler

Und dennoch: Der Gedanke, sich mit mehr als nur einer Disziplin auskennen zu sollen, mehr als nur eine einzige Sprache zu sprechen, wirkt auf manche belastend. Wie es die Sehnsucht nach einer normierten Einheitssprache gibt, so auch die Sehnsucht nach einer einzigen grundlegenden wissenschaftlichen Disziplin. Die Physiker sind bekanntlich der Meinung, dass eine solche grundlegende Disziplin nicht erst noch erfunden werden muss, sondern schon existiert, nämlich in Gestalt der Physik. Die Chemie ist dann weiter nichts als die Physik der äußeren Elektronenhülle. Und der Physiker von Weizsäcker sagte, dass „im Prinzip“ auch die Brüllaffenfamilie im Wald eine Lösung der Schrödingergleichung sei¹⁷. Das ist vermutlich richtig. Auch die Explosion der Atombombe über Hiroshima am 06.08.1945 dürfte eine Lösung der Schrödingergleichung gewesen sein, schließlich war diese Bombe ja von Physikern konstruiert worden. Aber einen rechten Einblick in die Tragweite des Ereignisses kann uns eben noch so viel Mathematik und Physik nicht bieten. Hier sind ganz andere Arten von Einsicht gefragt.

Und damit zeigt sich auch, dass die Physik zwar eine sehr vielfältige und bewundernswerte, aber keinesfalls eine umfassende Wissenschaft ist, die gewissermaßen auf Augenhöhe mit dem Sein der Dinge steht. Tatsächlich ist es eine von fast allen modernen Erkenntnistheoretikern geteilte Einsicht, dass alle menschliche Erkenntnis perspektivisch ist. Die ganze Wahrheit hat keine der heute tätigen Fachdisziplinen.

Deshalb ist es mehr als ein Luxus, wenn die Universitäten die Mehrsprachigkeit ihrer Mitglieder fördern. Und nun möchte ich noch einen weiteren Nutzen aus meiner Analogie ziehen, indem ich nämlich den Begriff der mehrsprachigen bzw. interdisziplinären Persönlichkeit einführe. In der Debatte um Interdisziplinarität wird sehr oft von Forschungsdesigns, Kooperationsformen, Institutionalisierungen und so weiter gesprochen. Dabei kommt es meiner Erfahrung nach ganz entscheidend auf die interdisziplinären Persönlichkeiten an, auf mehrsprachige Forscherinnen und Forscher, die gewissermaßen in zwei oder sogar drei Communities aufgewachsen sind, die zwischen den Gruppen vermitteln können, die interdisziplinäre Kooperationen moderieren können, ohne die alle rein institutionelle Förderung leicht

¹⁷ Weizsäcker (1986), S. 628.

leer läuft. Solche mehrsprachigen Persönlichkeiten, solche Doppelbegabungen sollten viel stärker gefördert werden, als es derzeit geschieht.

Interdisziplinarität ist zudem, auch daran erinnert die Analogie, nicht nur etwas, das sich herbeifinanzieren oder herbeiorganisieren lässt, sondern bedarf auch bestimmter kultureller Werte und Voraussetzungen, wie auch Mehrsprachigkeit durch bestimmte kulturelle und politische Gegebenheiten gefördert oder auch gefährdet werden kann. Intoleranz und ungleiche Machtverteilung begünstigen immer Einsprachigkeit, zur Mehrsprachigkeit gehören Toleranz, Fairness, Gleichberechtigung und zudem auch Offenheit gegenüber Neuem und Fremden. Eine entwickelte Kultur der Gastfreundschaft und der Philoxenie, der Offenheit gegenüber Menschen und Sichtweisen, die nicht dem angestammten Kollektiv angehören, ist aus meiner Sicht für wahre Interdisziplinarität ebenso unerlässlich wie für Mehrsprachigkeit.

6. Warum überhaupt Interdisziplinarität?

Zwei Motive lassen es heute wichtiger denn je erscheinen, Interdisziplinarität zu fördern, ein inneres und ein äußeres, ein wissenschaftsinternes und ein wissenschaftsexternes.

Der innere Grund ergibt sich aus den Veränderungen des Wissenschaftssystems. Dies war schon Helmut Schelskys Motiv für die Gründung des Zentrums für interdisziplinäre Forschung in Bielefeld. In einer Denkschrift formulierte er Anfang der 1970er Jahre: „die Re-Integration der sich spezialisierenden Wissenschaften zu einer Zusammenarbeit der verschiedenen Disziplinen“ sei eine zentrale Aufgabe, so schrieb er¹⁸. Heute muss man diese Forderung noch verstärken, denn die innere Ausdifferenzierung der Wissenschaft schreitet fort, und ohne Gegengewichte, die für eine innere Verbindung sorgen, die dafür sorgt, dass die in der Wissenschaft Tätigen auch das Bewusstsein haben, sich ein und derselben Institution zugehörig zu fühlen, ist die Wahrscheinlichkeit hoch, dass wir bald nicht mehr „die Wissenschaft“ haben, sondern nur noch ganz unterschiedliche Wissensprojekte, die einander wenig oder nichts zu sagen haben. Das Zentrum entfällt dann, die Wissenschaft wird dezentralisiert und reprivatisiert. Dies wird uns zwar dann von den Anstrengungen, einander zu verstehen und die Sprache der Anderen zu lernen, entheben,

¹⁸ Schelsky (1967), S. 72.

aber es wird auch dazu führen, dass die Wissenschaft der Öffentlichkeit gegenüber nicht mehr als Einheit wird auftreten können. Es wird dann nicht mehr dieses oder jenes Forschungsprojekt mit der Absicht gefördert, damit „die Wissenschaft zu fördern“. Vielmehr wird es darum gehen, dass diese oder jene Leute ganz gern etwas wissen wollen und um die Frage, ob das den Geldgebern die Summe wert ist oder eben nicht. Und was die Ergebnisse angeht, so wird es nicht mehr heißen: Die Wissenschaft hat herausgefunden, sondern: ‚Die und die Leute sagen, dass ...‘ Es gibt dann nicht mehr die Formulierung: ‚man weiß‘, sondern nur noch: ‚von der und der Institution wird behauptet, dass ...‘ Dass solche zentrifugalen Tendenzen heute schon unübersehbar sind, wird man kaum bestreiten können. Wenn wir also die Wissenschaft erhalten wollen, werden wir um die Bemühung, einander zu verstehen und miteinander zu kooperieren, nicht herumkommen. Hier setzt die Institutionenpolitik ein, und ich erinnere erneut an Schelsky, der keine philosophische Universalsynthese wollte, sondern „das gegenseitige begriffliche und methodische Verständnis der Disziplinen und die Entwicklung übergreifender theoretischer Konzeptionen“ fördern wollte¹⁹.

Aber um das zu erreichen, müssen, und da gehe ich über Schelsky hinaus, nicht nur Institutionen, errichtet werden, es müssen auch mehrsprachige Forscherpersönlichkeiten gefördert werden, hierfür müssen eigene Programme aufgelegt werden, für sie braucht es Entfaltungsmöglichkeiten und meiner Meinung nach auch eigene Lehrstühle.

Der *zweite Grund* für die Notwendigkeit, Mehrsprachigkeit mit drastisch gesteigerten Mitteln und neuen Instrumenten zu fördern, liegt in der Veränderung der Welt selbst.

Wenn es stimmt, dass wir im Anthropozän leben – der verstorbene Zoologe und DFG-bzw. Max Planck Chef Hubert Markl sprach, damals noch weitgehend ohne Resonanz, schon 1986 vom Anthropozoikum²⁰ – dann werden wir uns noch viel intensiver um Mehrsprachigkeit in den Wissenschaften bemühen müssen als bislang. Jedenfalls dann, wenn wir den Anspruch der Wissenschaft, zu erklären und vorherzusehen aufrechterhalten wollen. Oft stellt man sich ja die Zusammenarbeit von Natur – und Geistes- und Sozialwissenschaften so vor, dass die Naturwissenschaften die Probleme entdecken, deren Ursachen analysieren und daraus auch die

¹⁹ Ebd., S. 72f.

²⁰ Markl (1986), S. 324.

Lösung ableiten, während die Sozial- und Geisteswissenschaften dann ins Boot geholt werden, wenn die Lösung der sogenannten Öffentlichkeit vermittelt werden soll.

Eine solche Vorstellung ist aber absurd, wenn auch nur irgendetwas an der Vorstellung von einem Anthropozän dran ist. Im Anthropozän ist die Geistes- und Sozialwissenschaft nicht erst bei der Erarbeitung von Lösungen gefragt. Vielmehr muss auch die Ursachendiagnose interdisziplinär stattfinden! Und erst daraus kann man dann realistische Lösungsansätze herausbilden. Es gibt doch heute kaum noch eine Tatsache in unserer Biosphäre, die sich rein naturwissenschaftlich erklären ließe. Dass es Buckelwale gibt, z. B., ist nicht mehr nur ein Faktum, das sich rein evolutionstheoretisch erläutern ließe²¹. Es gibt sie vielmehr deshalb, weil die ansonsten sichere Ausrottung durch internationale Abkommen bislang verhindert wurde. Und diese Abkommen kamen zustande, weil sich aufgrund kultureller Prozesse in Europa und Amerika (Kanada, USA) das Bild des Buckelwals vom Monster hin zu einem musikalischen Geschöpf, zu einem *human of the sea* verändert hat. Hier erklären also kulturelle Phänomene und Prozesse das Vorliegen eines vermeintlich rein naturwissenschaftlichen Faktums. Und dies ist heute keineswegs nur noch eine Ausnahme, denn was von diesen Riesen gilt, gilt auch auf der Ebene der kleinsten Dinge.

Der Klimawandel, der Biodiversitätsschwund, die Austrocknung des Aralsees, die Austrocknung des Urmia-Sees, das Absterben der Korallen im Great Barrier Reef, die Toten Zonen im Golf von Mexiko oder auch moderne Epidemien wie Aids oder Covid-19 haben keineswegs rein natürliche Ursachen. Diese Phänomene des Anthropozäns lassen sich alle nur befriedigend erklären, wenn wir soziale Fakten einbeziehen. Diese sozialen Fakten sind, und das unterscheidet unsere Welt von der Welt vor 300 Jahren, als die moderne Naturwissenschaft entstand, unerlässlich für das kausale Verständnis sehr vieler ökologischer, klimatologischer und selbst geologischer Zusammenhänge. In den *Environmental Humanities*, aber zuvor auch schon in der sozialökologischen Forschung oder auch in der Politischen Ökologie und in etlichen weiteren Initiativen ist diese Einsicht grundlegend.

Sie ist es umso mehr, wenn man sich klarmacht, dass eine umfassende Ursachen-diagnose Voraussetzung für die Prognose ist. Betrachten wir ganz kurz auch hier ein Beispiel: Wie wird sich der Klimawandel weiter entfalten? Betrachtet man dies

²¹ Siehe Verf. (2019), S. 116–123.

allein durch eine naturwissenschaftlich-technische Brille, wird man kaum zu realitätsnahen Abschätzungen kommen, vielmehr wird man davon ausgehen, dass sich die Vernunft durchsetzt und die Verbrennung fossiler Energieträger innerhalb weniger Jahre gestoppt wird, weil es rational ist, die enormen Ausgaben, die der Klimawandel verursachen wird, zu mindern durch vergleichsweise geringe Ausgaben, die seine effiziente Bekämpfung heute verursacht²².

Wenn man aber neben den naturwissenschaftlich-technischen Realitäten auch politische und sozialwissenschaftliche einbezieht, kommt man vermutlich zu skeptischeren, wahrscheinlich auch realitätsnäheren Abschätzungen. Es ist nun einmal so, dass es ein internationales politisches System gibt, das anarchisch organisiert ist, nicht hierarchisch wie ein Staat. Es gibt nur eine Welt von Staaten, aber keinen Weltstaat. Und diese Staaten treffen ihre Entscheidungen auf der Grundlage ihrer Geschichte, nicht auf der Grundlage einer ahistorischen Vernunft. Die politischen Vertreter Chinas etwa denken nicht nur: Oh, der Klimawandel ist ein Problem, tun wir also alles dafür, es zu lösen. Man denkt in China auch: Wir lassen nicht erneut zu, dass andere Mächte unsere wirtschaftliche Entwicklung aufhalten. Und zudem sagt man: Das Problem wurde in erster Linie von anderen verursacht, von Großbritannien, den USA, auch von Deutschland, und das sind die Länder, die uns jetzt daran hindern wollen, uns unsererseits zu entwickeln²³. Mit anderen Worten: Der Klimawandel muss vor dem Hintergrund des Postkolonialismus gesehen werden. Diese Einsicht mag unbequem sein, unbequemer als die Annahme einer weltweit gleichen Vernunft, die auf Grundlage bereinigter Fakten das für ‚die Menschheit‘ Beste umsetzt. Aber sie eröffnet auch den Weg zu realistischen Abschätzungen und Strategien.

Wissenschaft verliert ihre Hellsicht und ihre Weitsicht, auch ihre praktische Brauchbarkeit, wenn sie im Umweltbereich auf konsequente Interdisziplinarität verzichtet. Sie kann weder ausreichend erklären noch ausreichend vorhersehen, sie kann die Gesellschaft nicht orientieren, wenn sie sich nur auf einzelne Disziplinen stützt. Und damit kommen wir auf Lichtenberg zurück. Wer nur naturwissenschaftlich forscht, kann heute die Natur nicht mehr verstehen. Fördern wir also die Mehrsprachigkeit! Sie wird auch eine neue Lust am Forschen, eine neue Lust am Verstehen und Lernen sein.

²² Siehe Verf. (2020).

²³ Vgl. Gebauer (2017), insbesondere S. 380.

Literatur

- Bachelard, Gaston (1972): *La formation de l'esprit scientifique*, Paris
- Gebauer, Sebastian (2017): China und andere Schwellenländer, in: Georg Simonins, Hg., *Handbuch Globale Klimapolitik*, Paderborn, S. 377–396
- Jungert, Michael u. a., Hg. (2013): *Interdisziplinarität. Theorie, Praxis, Probleme*, Darmstadt
- Köller, Wilhelm (2006): *Narrative Formen der Sprachreflexion*, Berlin
- (2012): *Sinnbilder für Sprache. Metaphorische Alternativen zur begrifflichen Erschließung von Sprache*, Berlin–Boston
- Krohn, Wolfgang/Grunwald, Armin/Ukowitz, Martina (2017): Transdisziplinäre Forschung revisited: Erkenntnisinteresse, Forschungsgegenstände, Wissensform und Methodologie, in: *GAIA – Ecological Perspectives on Science and Society* 26/4, S. 341–347
- Lichtenberg, Georg C. (1983): *Sudelbücher. Hg. von Franz H. Mautner*, Frankfurt am Main
- Luhmann, Niklas (1990): *Die Wissenschaft der Gesellschaft*, Frankfurt am Main
- Markl, Hubert (1986): *Natur als Kulturaufgabe. Über die Beziehung des Menschen zur lebendigen Natur*, Stuttgart
- Mittelstraß, Jürgen (1989): *Der Flug der Eule. Von der Vernunft der Wissenschaft und der Aufgabe der Philosophie*, Frankfurt am Main
- (1994): Hochschulkultur. Die Anstrengung des Begriffs und die Lust des Studierens, in: ders., *Die unzeitgemäße Universität*, Frankfurt am Main, S. 88–100
- (2018): Forschung und Gesellschaft: Von theoretischer und praktischer Transdisziplinarität, in: *GAIA – Ecological Perspectives on Science and Society* 27/2, S. 201–204
- Neumann, Gerhard (1976): *Ideen-Paradiese. Untersuchungen zur Aphoristik von Lichtenberg, Novalis, Friedrich Schlegel und Goethe*, München
- Schelsky, Helmut (1967): Das Zentrum für interdisziplinäre Forschung. Eine Denkschrift, in: Paul Mikat/Helmut Schelsky, *Grundzüge einer neuen Universität. Zur Planung einer Hochschule in Ostwestfalen*, Gütersloh, S. 71–87
- Soentgen, Jens (2014): Volk ohne Stoff. Vom Mythos der Ressourcenknappheit, in: *Merkur. Deutsche Zeitschrift für europäisches Denken* 68/777, S. 182–186
- (2019): *Ökologie der Angst*, Berlin
- (2020): Am Ende des Zwei-Grad-Zieles. Für einen Strategiewechsel in der Klimapolitik, in: *Merkur. Deutsche Zeitschrift für europäisches Denken* 74/849, S. 22–33
- (2021): Umweltforschung im Anthropozän. Warum Interdisziplinarität unerlässlich ist, in: *Merkur. Deutsche Zeitschrift für europäisches Denken* 75/860, S. 89–95
- Voigt, Uwe (2013): Interdisziplinarität: Ein Modell der Modelle, in: Michael Jungert u. a. (2013), S. 31–46
- Wandruszka, Mario (1979): *Die Mehrsprachigkeit des Menschen*, München–Zürich
- Weizsäcker, Carl F. von (1986): *Der Aufbau der Physik*, München

Facetten natürlicher Intelligenz am Beispiel des Brown Bear/Grizzlybär (*Ursus arctos horribilis*)

Annette Belke

Zusammenfassung

Gegenwärtig ist es nicht möglich, die Intelligenz von Grizzlybären direkt zu messen und mit dem Intelligenzquotienten anderer Säugetiere zu vergleichen, denn ein artübergreifendes Testsystem steht noch aus. Die Reflexion über strategisches und zielgerichtetes Verhalten von Grizzlybären beruht daher auf Beobachtungen und Anekdoten, wie offenbar auch David Humes grundlegende Annahme, der zufolge Tiere ebenso wie Menschen mit Denken und Vernunft ausgestattet sind.

Im Fokus sind die spezifischen Fähigkeiten des Grizzlybären. Allen voran steht eine Beschreibung seiner neuromuskulären Geschicklichkeit, seiner hochdifferenzierten Sinne, seines Sozialverhaltens und seines genetisch gesteuerten Verhaltens im Winter. Auch mag sich die Frage erheben: Können Grizzlybären lächeln und lachen?

Der Blick nach vorne richtet sich auf die Zukunft von *Arctos Ursus* im Anthropozän. Obwohl die Grizzlybärpopulationen in Nordamerika rasant zurückgehen, werden diese Bären nach wie vor gejagt – meist aus Prestige Gründen, um dem Ego von Leuten wie Donald Trump Jr. bei seinem jüngsten Aufenthalt im kanadischen Yukonterritorium zu schmeicheln.

Grizzlybären sind hochintelligente, soziale Säugetiere. Die Aufzucht ihrer Jungen ist äußerst zeitintensiv. Was gibt uns das Recht dazu, sie zu töten? Gemäß den Traditionen der *First Nations* heißen Bären „Großmutter“ oder „Großvater“ und verdienen Respekt als „die anderen Menschen“.

Abstract

Currently, the intelligence of grizzly bears cannot be measured directly and compared with the intelligence quotient of other mammals, as the development of an inter-specific test system is pending. The reflection on the possible strategic and targeted behavior of grizzly bears is therefore based on observations and anecdotes, which seem to have substantiated David Hume's basic assumption: "Beasts are endowed with thought and reason as well as men".

A sketch of the grizzly's specific abilities is attempted, describing his neuromuscular agility, his highly differentiated senses, his social behavior, his genetically controlled winter behavior and the question may be posed: can grizzly bears smile and laugh?

The outlook points to the future of *Arctos Ursus* in the Anthropocene. In the time of rapidly declining grizzly bear populations in North America these bears are still hunted for their prestige value, to assuage the egos of people like Donald Trump Jr. on his recent swing through the Yukon Territory, Canada.

Bears are a highly intelligent, social species, who place great store in rearing their young. What gives us the right to kill them? First nation traditions call them "Grandfather, Grandmother", and respect them as "the other humans".

1. Vorannahmen

Wir unterscheiden, definieren und klären: Was ist natürliche und was ist künstliche Intelligenz, von der wir erwarten dürfen, die Ursachen unserer Umweltprobleme zu verstehen und diese Probleme hoffentlich im verzahnten Miteinander beider Intelligenzformen lösen zu können.

Genauer betrachtet gibt es keinen Gegensatz zwischen den beiden Intelligenzformen, da die eine in der anderen enthalten ist – wenn denn das Natürliche dasjenige ist, das wir benennen können, ohne es schon begriffen haben zu müssen. Wenn wir dementsprechend unter dem Künstlichen dasjenige verstehen, was wir umgekehrt (gestaltend) begreifen können, ohne es schon benannt haben zu müssen, dann zeigt es sich, dass Intelligenz unter je verschiedener Hinsicht natürlich und künstlich zugleich sein kann.

Unsere „natürliche“ Intelligenz ist auch künstlich, insofern sie kulturell geformt ist; künstliche Intelligenz ist als zunehmend autonomer Teil unserer Umwelt für uns auch natürlich und existiert in einer eigenen Umwelt, in der ihr für sie Natürliches begegnet¹.

Der Grizzlybär ist Kanadas Ikone für unberührte Wildnis. Doch auch vor dem *Ursus arctos* macht das Anthropozän nicht Halt. Diese Spezies gilt selbst im Hohen Norden als *endangered species* oder *species at risk*. Das Überleben der *Brown Bears* ist unter dem zunehmenden Zuzug und Zugriff des Menschen erheblich gefährdet. In einigen kanadischen Provinzen gelten Grizzlybären bereits als *extirpated*, ausgerottet, z. B. in Alberta. Der Umgang mit dieser Spezies veranschaulicht die Auswirkungen des Anthropozäns und legt die Notwendigkeit nahe, die Ressourcen natürlicher und künstlicher Intelligenz zu bündeln, um dieser Entwicklung entgegen zu handeln.

2. Komplexität und Intelligenz

Der Grizzlybär ist genetisch, motorisch intelligent und entfaltet Problemlösungsstrategien je nach Lebensraum und den damit verbundenen Herausforderungen. Im Folgenden betrachten wir die Komplexität des Bären als Beispiel für das Ineinander von natürlicher und künstlicher Intelligenz im oben definierten Sinn.

3. Biologische Nomenklatur

Im Gebrauch ist hier die englischsprachige Unterscheidung, da die von mir verwendete Literatur sich auf dieselbe bezieht. Die biologische Art *Brown Bear* (*Ursus arctos*) wird unterteilt in zwei Unterarten:

- a. *Kodiak Bear* (*Ursus arctos middendorffi*), auch *Alaskan Brown Bear* genannt, und
- b. *Grizzly Bear* (*Ursus arctos horribilis*).

¹ Vgl. dazu U. Voigt (2021).

4. *First Nations* und „Großvater“

Indigene Bewohner, in Kanada allgemein *First Nations* genannt, verehr(t)en Bären als eine Art Über-Mensch oder zumindest betrachteten sie den *Ursus arctos* als menschenähnlich². Sie titulierte *Brown Bears* häufig ehrfurchtsvoll als „Großvater oder Großmutter“.

In der ersten bekannten schriftlichen Notiz zum Thema Grizzlybären bemerkt der Jäger Henry Kelsey im Jahr 1691, dass die „native Indian tribes“ diese Tiere als Götter verehrten³. Leider sind viele indigene Traditionen nur noch in Fragmenten erhalten oder abgebrochen, da sie ausschließlich mündlich tradiert wurden.

Bären wurden aufgrund ihrer außergewöhnlichen Fähigkeiten respektiert und gleichsam als übernatürliche Verwandte betrachtet. Noch weitgehend intakte Überlieferungen kanadischer Ureinwohner wie die der Tlingits bezeichnen Grizzlybären als „the other humans“⁴.

Phil Timpany, *Grizzly Bear Guide* und Fotograf, nimmt an, dass Bären Menschen als ihresgleichen einstufen, „da wir auf zwei Beinen gehen wie sie es auch zu tun vermögen. [...] Bären nehmen den *Homo erectus* womöglich als eine mit anderen Fähigkeiten begabte Bären-Art wahr“⁵. Eine wilde Spekulation oder ein Körnchen Wahrheit?

5. Der Körper des Bären

5.1. Gehirnvolumen

Der zerebrale Kortex von Grizzlybären ist stark ausgeformt, verglichen mit dem gesamten Gehirnvolumen. 45 % des Gehirns eines Grizzlybären besteht aus dem zerebralen Kortex. Der Kortex ummantelt das Zerebrum, das den größten und komplexesten Teil des Gehirns darstellt. Dieser ist verantwortlich für die Entstehung und Verarbeitung von Sinneswahrnehmungen und steuert die motorischen und assoziativen Funktionen.

² Phil Timpany, Persönliches Gespräch im Takku River Grizzly Bear Camp (2019).

³ Busch (2000), S. 104.

⁴ Eaton (2020).

⁵ Phil Timpany (2019), a. a. O.

Der zerebrale Kortex einer Maus beinhaltet nur 30 % des gesamten Gehirnvolumens dieser Nagerart. Bei Schimpansen wurde das Verhältnis Kortex/Gehirnvolumen auf etwa 55 % festgelegt⁶.

Das Zerebrum, bestehend aus rechter und linker Gehirnhälfte, bildet den größten Teil des Gehirns. Von dort werden Funktionen gesteuert wie die Interpretation von haptischen Eindrücken, von visuellem und auditivem Input, aber auch Sprachfähigkeit, schlussfolgerndes Denken, Emotionskontrolle, Lernfähigkeit und Feinmotorik.

Der zerebrale Kortex ist aktiv im Zusammenhang mit Handlungen, die gestaltungsoffen sind, also einen gewissen Spielraum beinhalten. In diese Kategorie fallen das Erlernen der Muttersprache und von Fremdsprachen und die Art und Weise sich auszudrücken sowie verschiedene andere Funktionen des Gehirns, genannt seien hier stellvertretend Denken und Erinnerungsfähigkeit. Kennzeichnend für den Kortex ist dessen Komplexität und Verbundenheit mit anderen Teilen des Gehirns.

Derzeit kann die Intelligenz von Bären noch nicht direkt gemessen und mit dem Intelligenzquotienten anderer Säugetiere verglichen werden. Deshalb basiert die Reflexion über mögliches strategisches und oder zielgerichtetes Verhalten von Bären auf Beobachtungen und Anekdoten.

Viele intelligente Handlungen von Grizzlybären wurden beobachtet und beschrieben. Beispielsweise entschärften Grizzlybären Fallen und verzehrten anschließend das in ihnen enthaltene Lockmittel. Wir können annehmen, dass sie die Falle öffneten, um an deren Inhalt zu gelangen oder sie sich nur aus Neugierde damit beschäftigt und zufällig den Köder entdeckt haben. Ein weiteres Beispiel: Eine Grizzlybär Mutter grub mit den Vorderpfoten Tritte/Stufen für ihre Jungen in einen für diese schwer zu überwindenden Bergrücken⁷.

James Capen, alias *Grizzly Adams*, trainierte gefangene Grizzlies. Mit hohen Worten lobt er seine Schüler: „Grizzly Bear behaviour may be modified and improved to such a degree as to be a wonder“⁸. Doug Seus, ein Tiertrainer aus Utah, bildete Grizzly- und Kodiakbären als „Filmschauspieler“ aus. Bekannt wurde sein 682 kg schwerer Schüler, Bart, ein Kodiakbär, der die entsprechen-

⁶ Busch (2000), S. 90–92.

⁷ Ebd., S. 94.

⁸ Olsen (1990), S. 3.

den Posen einzunehmen wusste. Seus: „My Grizzlies and Kodiaks are the hardest to tame, but the easiest to train, generally you only have to teach them something once.“⁹

5.2. Physische Fähigkeiten

Grizzlybären sehen behäbig aus, aber auf Kurzstrecken sind sie so schnell wie ein durchschnittliches Pferd. Für etwa zehn Kilometer vermögen sie eine mittlere Geschwindigkeit zu halten. Sie sind also Sprinters und Ausdauerläufer. Grizzlybären sind exzellente Schwimmer. Ein Grizzlybär ist etwa so stark wie drei bis fünf Menschen, seine Stärke nimmt zu, wenn er oder sie in Rage gerät¹⁰. Mit Hilfe seiner starken Kiefern kann der Grizzlybär Knochen und Schädel zermalmen. In Erstaunen versetzt seine Fähigkeit, schnell und leicht Bäume zu erklettern. Auch in schwierigem Gelände ist der Grizzlybär äußerst wendig.

5.3. Neuromuskuläre Beweglichkeit

Studien von Frank und John Craighead an Grizzlybären im Yellowstone Nationalpark¹¹ belegen, dass die Bären ihre Vorderbeine bzw. Arme um die Achse zu drehen vermögen.

Ihre Wendigkeit ist möglicherweise der von Primaten sogar überlegen. Es gibt noch keine vergleichenden Studien. Ein Grizzlybär kann die Krallen einzeln bewegen und damit behutsam Tannenzapfen öffnen oder einen gut verschlossenen Müll-eimer öffnen¹². Es wurde auch beobachtet, dass ein Grizzlybär einen Pfirsich mit Hilfe einer einzigen Kralle entkernt, die durch die Frucht durchgestochen wird¹³. Geschicklichkeit entsteht aus angemessenem, zielorientiertem Verhalten und ist damit eine sich in und mit der Physis verwirklichende Form von Intelligenz.

⁹ Busch (2000), S. 93.

¹⁰ Ballard (2018), S. 331.

¹¹ Craighead (1982).

¹² Ballard (2018), S. 330ff.

¹³ Olsen (1990), S. 4.

5.4. Sinneskanäle und Wahrnehmung

Der Grizzlybär besitzt einen hochdifferenzierten Geruchssinn. Die ausgeprägte Nasenhöhle enthält ein komplexes Netzwerk von nasalem Schleim, rund hundertmal so umfangreich wie das eines Menschen. Damit ist die Geruchsempfindlichkeit des Grizzlybären etwa tausendmal so scharf und ausgefeilt wie die eines Menschen. „If we were endowed with the sensory system of a Grizzly for a minute we would faint right away.“¹⁴

Grizzlybären riechen z. B. Aas aus einer Entfernung von mehreren Kilometern. Sie informieren sich zudem über einander durch Beriechen von Urinabsatz und Pfotenabdrücken. Ein Grizzlybär entnimmt aus diesen Anzeichen der Präsenz eines anderen Bären, ob es sich um ein Weibchen handelt, einen Rivalen oder ein subordiniertes Männchen. Der Grizzlybär vermag auch aus den obigen Anzeichen den Hormonzustand eines Weibchens zu lesen¹⁵.

Grizzlybären haben ein zusätzliches geruchserkennendes Organ, das sogenannte Jacobson-Organ. Mit dessen Hilfe decodiert der Grizzlybär Gerüche, die in Zusammenhang mit Feuchtigkeit entstehen. Das Organ sitzt am Gaumen. Es wird in der innerartlichen Kommunikation eingesetzt, um zu riechen, ob ein anderer männlicher Grizzlybär überlegen und daher zu vermeiden ist oder ob ein Grizzlybär-Weibchen paarungsbereit ist.

Der Sehsinn des Grizzlybären entspricht nach derzeitigem Wissensstand ungefähr dem des Menschen. Grizzlybären wurden oft dabei beobachtet, wie sie auf entfernte Stimuli reagierten. Ihr Gesichtsfeld und ihre Farbwahrnehmung scheint jeweils dem menschlichen Gegenstück gleichzukommen. Um noch genauer sehen zu können, stellt sich der Grizzlybär auf die Hinterbeine, was oft als aggressive Geste missverstanden wird. Zusätzliches Gewebe hinter den Augäpfeln wirft Licht zurück zur Retina. Bei Dämmerung oder in der Nacht wird so die Sehkraft verstärkt. Die Fähigkeit des Grizzlybären, bei Dunkelheit zu sehen, ist der des Menschen überlegen.

¹⁴ Phil Timpany, Persönliches Gespräch im Takku River Bear Camp, August 2019.

¹⁵ Ebd.

Die Ohren sind zwar klein im Verhältnis zum Körper, Experimente ergaben jedoch, dass der auditive Sinn des Grizzlybären etwa dem des Menschen entspricht¹⁶.

Tastsinn: Bei so viel Muskelkraft erstaunen die feinmotorischen Möglichkeiten des Grizzlybären, die bereits erwähnt worden sind

Es gibt derzeit noch keine wissenschaftlich eruierten Informationen zum Thema Geschmackssinn des Grizzlybären.

5.5. Körperimmanente Vorgänge biologischer Intelligenz

Entgegen allgemeiner Auffassung treten Bären keinen Winterschlaf an. Der Zustand ihres Körpers bei kalten Temperaturen wird von Seiten der Biologie als eine Form von Torpor oder Erstarrung eingeordnet.

„*Torpor* is a survival tactic used by animals to survive the winter months. It also involves a lower body temperature, breathing rate, heart rate, and metabolic rate“¹⁷. Aber anders als der Winterschlaf scheint der Torpor, im Deutschen mit „Winterruhe“ übersetzt, ein unfreiwilliger Zustand zu sein, der diktiert von den klimatischen Bedingungen die Grizzlybären überkommt. „But unlike *hibernation*, *torpor* appears to be an involuntary state that an animal enters into as the conditions dictate“¹⁸.

In der Zeit der Winterruhe fällt die Körpertemperatur der Grizzlybären auf etwa 30 bis 35 °C. Die Herzfrequenz der Bären verlangsamt sich auf 8 bis 9 Schläge in der Minute. Säugetiere, die in den Winterschlaf verfallen, haben jedoch eine wesentlich tiefere Herzfrequenz. Während der Winterruhe wachen Bären immer wieder auf, strecken oder bewegen sich.

Drastisch verwandeln sich in dieser Zeit ihre Körpergrundfunktionen. Bis zum Frühjahrsbeginn werden sie weder urinieren noch defäkieren. Die abzustößenden Feststoffe verwandeln sich in einen Pfropfen, der im Frühjahr abgesondert wird. Was besonders erstaunlich ist: Der Bärenkörper verwandelt Harnstoffe in Protein. Nierenspezialisten versuchen derzeit diese Vorgänge medizinisch zu erklären. Soll-

¹⁶ Vgl. <https://ssl.eas.ualberta.ca/safety/bear-safety-information-awareness-avoidance>, (Zugriff: 19.07.2021).

¹⁷ Grizzly Bear Foundation (2019), News, Grizzlybearfoundation.com.

¹⁸ Ebd.

te dies gelingen, ergeben sich neue Behandlungsmöglichkeiten für Menschen mit Nierenproblemen.

Wären Bären nicht dazu in der Lage, während der Winterruhe Harnstoffe in Protein zu verwandeln, würden diese bald ihren Körper von innen vergiften. Die außerordentliche Fähigkeit, Harn in Protein zu verwandeln, erklärt auch, warum Bären während der langen Zeit der Inaktivität keine Muskeln abbauen¹⁹.

Ein weiteres „Wunder“ ereignet sich während des Winterschlafs. Gegen Ende Januar bis Ende Februar werden die Jungen geboren. Da die Mutter sich während dieses Vorgangs noch in Winterruhe befindet, müssen die Kleinen, die nicht größer sind als Eichhörnchen, selbst herausfinden wie sie zu den Zitzen der Mutter gelangen und diese dazu animieren, Milch abzugeben. Die blind geborenen Jungen sind in der Regel einige Wochen auf sich gestellt, bis zum Frühjahrserwachen ihrer Mutter. Was wie ein gemütlicher, langer Schlaf erscheint, ist tatsächlich eine geniale Form der Überbrückung von härtesten Wetterbedingungen und der damit verbundenen Nahrungsknappheit²⁰.

6. Soziale Intelligenz

6.1. Hierarchiebildung und vokales Ausdrucksverhalten

Die soziale Intelligenz dieser Bären zeigt sich schon daran, dass sie Hierarchien bilden, etwa wenn sie sich an ertragreichen Futterplätzen treffen entwickelt sich eine soziale Hierarchie. „The bears know and remember each other as individuals, and have memories that span many years“²¹.

Zudem kommunizieren Grizzlybären miteinander unter Einsatz unterschiedlicher Laute und unter Verwendung einer differenzierten Körpersprache. Sie verfügen über eine Varianz von Lauten, die sie in der sozialen Interaktion einsetzen, wie z. B.: „Grunts of contentment, purring sounds, noises of distress like screaming, blowing sounds, grunting, barks, moaning, woofing sounds, loud agitated huffing“²².

¹⁹ Ebd.

²⁰ Ebd.

²¹ Olsen (1990), S. 4.

²² Ballard (2018), S. 333f.

Klare Geruchsnachrichten hinterlassen sie an den Kratz-Bäumen (*scratching posts*) und via Baum-Markierungen (*tree marks*). Der diese verursachende Grizzlybär berichtet so über seine Stärke und seinen Gesundheitszustand.

6.2. Interaktion

Die soziale Interaktion zwischen Erwachsenen und Jungbären ist sehr ausgefeilt. „As highly evolved social animals, bears form hierarchies and have structured relationships with each other, sometimes even sharing resources.“²³ Das klassische Bild des Grizzlybären und des Eisbären ist das eines einsamen Wanderers, der nur in der Paarungszeit Zweisamkeit sucht bzw. in der Familienphase beschränkt auf Mutter und Kinder soziale Bindungen pflegt. „In fact, the polar bear, typically thought of as solitary, actually lives within a community of other polar bears and never loses track of other members.“²⁴ Lange andauernde Bindung bis ins dritte Lebensjahr der Jungbären ist Voraussetzung für deren Überlebensfähigkeit²⁵.

6.3. Beispiele altruistischen Verhaltens

Mehrfach wurde beobachtet, dass verwaiste Junge von fremden Bärinnen adoptiert wurden. Bekannt und bestätigt ist ebenfalls, dass verletzte Grizzlybären von Artgenossen umsorgt werden²⁶.

Grizzlybären leben häufig miteinander in kleinen zusammengesetzten Gruppen. Vor diesem Hintergrund erklärt sich, dass Bären durchaus fähig sind, so-

²³ Morgan (2008), Bears of the Last Frontiers, in: Nature, www.pbs.org.

²⁴ Ebd.

²⁵ Ebd.: „By the time a juvenile bear leaves its mother, it knows what plant foods are available at each time of the season, and what habitats are likely to have those foods over a very large area. That knowledge serves them well as they move into new areas, learning and remembering where new food sources are found in a new environment. This knowledge is also critical to finding food when food sources change drastically from year to year depending upon weather and climate.“

²⁶ Busch (2000), S. 88–90.

ziales Verhalten zu zeigen. „Some quite sociable behaviour has been noted on occasion“²⁷.

6.4. Komplexes Paarungsverhalten

Das Werbungs- und Paarungsverhalten von Grizzlybären ist vielgestaltig und kreativ. Ihre „Courtship“ weist Elemente auf wie Arm umlegen und lang andauerndes Küssen²⁸.

7. Intelligente Gewohnheiten und Verhaltensweisen

Das *Kodiak Bear Center Alaska* siedelt die Intelligenz dieser Bären zwischen der eines hochintelligenten Hundes und der eines durchschnittlichen Primaten an. Wie beim Menschen variiert der Intelligenzgrad je nach Individuum. Ebenso besitzt jeder Bär eine eigene Persönlichkeit, abhängig von seinen Lebenserfahrungen und seinem genetischen Erbe. Als „opportunistic omnivore“ sind Bären dazu in der Lage, jede Art von Nahrung zu sich zu nehmen. Diese Voraussetzungen befähigen sie zu elaboriertem und komplexen Problemlösungsverhalten²⁹.

Die Gewohnheiten und Verhaltensweisen der Bären verweisen auf ihre Intelligenz. Sie erinnern sich an unterschiedliche Futterplätze. Sie merken sich wann, welche Früchte oder Beeren reif und wo diese jeweils zu finden sind. Ein Grizzlybär kann sich an Futterplätze erinnern, die er vor zehn Jahren oder früher zum letzten Mal aufgesucht hat. Bären erinnern sich an andere Bären und können deren sozialen Status aus einer Distanz von 600 m Entfernung erkennen³⁰.

98 % des ursprünglichen Grizzlybär-Habitats ist unter menschlicher Kontrolle. Die Bären suchen deshalb zwangsläufig in unserem näheren und/oder weiterem Umfeld nach Essbarem. Oft bleibt ihnen angesichts der Futterknappheit keine andere Wahl. Aufgrund von Überfischung und dem anthropogen verstärkten Klima-

²⁷ Busch (2000), S. 87.

²⁸ Murie (1981) 88f.

²⁹ Kodiak bear center (2019), Brown Bear Facts, kodiakbearcenter.com.

³⁰ Morgan (2008), Bears of the Last Frontiers, in: Nature, www.pbs.org.

wandel geht die Lachswanderung drastisch zurück. Extreme Winter und Sommer Trockenheit lassen ihre natürliche Kost – die Kanadische Büffelbeere, Blaubeeren, Moosbeeren, Moltebeeren – nicht mehr reifen. Sogenannte *Bear attractants* wie Bienenstöcke, Hühner Häuser mit Ausläufen, usw. werden meist nicht mit elektrischen Zäunen gesichert.

Bären versuchen kreativ, an diese Ressourcen zu gelangen. Es gelingt ihnen, Mülleimer, Haus- und Autotüren zu öffnen. Abfälle von Campern veranlassen Grizzlybären dazu, auf variantenreiche Art und Weise Müll-Container in Nationalparks zu öffnen. Bei dem Versuch, an das begehrte Gut zu gelangen, testen sie die jeweils neue Generation der so genannten *Bear Proof Container*. Neue Modelle werden mit Hilfe von Grizzlybären in Gefangenschaft auf ihre Tauglichkeit untersucht³¹.

8. Selbstgewahrsein

Was ein natürliches Wesen von einem Roboter unterscheidet, ist u. a. eine gewisse Form von Selbstwahrnehmung. Grizzlybären-Jäger William Wright bemerkte in seinem Buch *The Grizzly Bear: The Narrative of a Hunter-Naturalist* (1909), er habe erlebt und beobachtet, dass Grizzlybären sehr geschickt darin seien, ihn zu vermeiden: „In the beginning, I studied the grizzly bear in order to hunt him. I came to hunt him in order to study him. I laid aside my rifle.“³²

1977 fasste der Grizzlybär-Forscher Frank Craighead im Vorwort zur überarbeiteten Neuauflage von Wrights Werk dessen Grizzly-Beobachtungen grundlegend zusammen. Wrights Kommentare berichten durchgehend von der Durchtriebenheit (*cunning*) und Bauernschläue (*shrewdness*) des Grizzlybären. Sicher handelt es sich um Anthropomorphismen, die allerdings auf die Komplexität des Grizzlybären hinweisen³³. Wrights Berichte, in denen er gerne den Begriff *self-concealment* benutzt, decken sich mit den Beobachtungen neuer Grizzlybär-Forscher, die übereinstimmend behaupten, die Tiere seien dazu in der Lage, sich zu verstecken.

³¹ Ebd.

³² Olsen (1990), S. 3.

³³ Ebd.

Lance Olsen³⁴, University Montana, zitiert in seinem Artikel „Cognitive Complexity of the Grizzly Bear“ einen weiteren frühen Naturforscher: Enos Abija Mills. Dieser schrieb 1919 in *The Grizzly, our greatest wild animal*, dass Grizzlybären sich verstecken und ihre Verfolger im Auge behalten. „All Grizzlies are Scouts of the first order; they are ever on guard.“³⁵ Olsen beschreibt, dass die Bären dazu in der Lage seien, Verstecke zu finden, die es ihnen ermöglichten, ihre Umgebung zu beobachten, ohne dabei selbst gesehen zu werden.

Olsen kommt zu dem Schluss, dass Grizzlybären über eine Art von Selbstbewusstsein verfügen, da er annimmt, dass dies eine Voraussetzung dafür ist, sich erfolgreich verstecken zu können. Es mag anthropomorph klingen, den Bären zu unterstellen, dass sie absichtlich – wie oft berichtet – menschliche Verfolger in die Irre führen, sie von Wegen abbringen und deren Orientierungssinn verwirren. Aber vorab anzunehmen, sie seien nicht dazu in der Lage, könnte auf Anthropozentrismus hinweisen.

9. Und sie lächeln?

Dieses Kapitel greift zurück auf Gedanken neuzeitlicher Philosophie und tastet sich heran an mögliche kognitive Forschungsprojekte der Biologie, spezifisch an Grizzlybären. René Descartes lange unumstrittene Behauptung, Tiere besäßen keine Intelligenz, sondern nur Instinkt, ist längst widerlegt³⁶. Descartes spricht aufgrund seines mechanistischen Ansatzes³⁷ unseren tierischen Verwandten darüber hinaus sogar jegliche Empfindung ab. Unterscheidendes Merkmal von Menschen und Tieren ist für den französischen Philosophen, dass Tiere keine *mens* (Englisch: *mind*) besitzen, hier wohl am besten mit Vernunft zu übersetzen, und damit auch keine Sprache. „In other words, hitting a dog with a stick, for example, is a kind of input and the squeal that follows would be merely output, but the dog did not feel anything at all and could not feel pain unless it was endowed with a mind“³⁸.

³⁴ Ebd.

³⁵ Ebd.

³⁶ Regan/Singer (1989), S. 13–19.

³⁷ Ebd.

³⁸ Vgl. Internet Encyclopedia of Philosophy, Descartes, 8d, iep.utm.edu, Zugriff: 14.01.2021.

Im Verständnis der Tiere folgen wir heute weitgehend dem Empiriker David Hume: „No truth appears to me more evident than that beasts are endowed with thought and reason as well as men.“³⁹ Die Frage, ob Tiere leiden und Schmerzen empfinden, die der Philosoph Jeremy Bentham mit einem eindeutigen „Ja“ beantwortete, bildet den philosophischen Hintergrund der Überlegung, ob Tiere sich zu freuen vermögen. „The question is not, Can they reason?, nor Can they talk? but, Can they suffer? Why should the law refuse its protection to any sensitive being?“⁴⁰

Wenn wir anerkennen, dass Tiere die Fähigkeit besitzen, Leid und Freude vielgestaltig zu erleben, erfolgt eine weitere Frage: Sind Tiere in der Lage zu lächeln? Schmerz Ausdruck ist uns von vielen Säugetieren bekannt, während das Phänomen des Lächelns als Ausdruck der Freude eine delikate Frage westlicher Philosophie berührt. Sind das Lachen und seine milde Ausdrucksform, das Lächeln, nicht nur dem Menschen vorbehalten⁴¹? Davon überzeugt ist jedenfalls bereits Aristoteles. Von seinen Schriften ausgehend tanzt diese Auffassung durch das Mittelalter bis in die frühe Neuzeit.

Eine massive Gegenstimme findet sich endlich bei Charles Darwin: „A pleasurable and excited state of mind, associated with affection, is exhibited by some dogs in a very peculiar manner; namely by grinning.“⁴² Die Hundeforscherin Feddersen-Petersen führte Langzeit-Studien mit Hunden durch, die genau dieses Ergebnis belegen. Hunde sind in der Lage zu lächeln; rassebedingt mehr oder weniger sichtbar, je nach den vorhandenen mimischen Fähigkeiten kann sich entspannte Freude im Gesicht von Hunden zeigen. Die Biologin unterscheidet zwischen zwei Typen des „Lächelns“, „eine Form tritt nur bei Hunden im submissiven Kontext sowie zur Kontaktaufnahme mit dem Menschen auf, die andere als sehr entspannter Ausdruck im Sozialspiel“.⁴³

Lächeln ist beim Menschen genetisch angelegt⁴⁴. Bei Erwachsenen steht die Fähigkeit im Zusammenhang von sozialer Zuwendung, verweist aber auch auf Selbst-

³⁹ Hume.

⁴⁰ Bentham (1789), in: Animal Rights Philosophy, speakingofresearch.com, Zugriff: 14.01.2021.

⁴¹ Vgl. Aristoteles, De partibus animalium III 10, 673a9.

⁴² Darwin (1874), The Expression of Emotions in Man and Animals, 3rd ed., in: Ekman, P. (Ed.), Harper Collins, London.

⁴³ Feddersen-Petersen (2008), S. 131.

⁴⁴ Ebd.: „Lächeln ist eine angeborene Verhaltensweise, ein Ausdruck der freudigen Zuwendung“. Lächeln erweist sich „reaktiv auf bestimmte Stimulationen als unabhängig von der

reflexion und innwendige Kommunikation mit vorgestellten Partnern oder Partnerinnen.

Wenn wir die oben erwähnten Beobachtungen und Ergebnisse der Hundeforschung auf die Erforschung der Innen- und Außenwelt von Grizzlybären übertragen, stoßen wir auf die emphatischen Worte der Forscherin Poulsen, die deren Nachruf in *National Geographic* überliefert:

„Bears smile just like we do. They pull each side of their mouth upwards. They smile for their reasons of self-contentment, just like we humans smile for our reasons of self-contentment. It's just that our reasons may not be the same.“⁴⁵

Die Emotionen, die Bären körpersprachlich zeigen, beziehen sich auf das, was in ihrer Bären-Welt zählt. Beispielsweise mag ein Grizzlybär in Montana lächeln, wenn er einen Berggipfel erreicht und dort Tausende für ihn essbare Larven findet. Für uns höchstens ein Grund dazu, angeekelt wegzusehen.

„We as a society have been slow to allow animals to have the converse feelings of love, happiness, joy – those kinds of things. We think only humans can have those feelings.“⁴⁶

In gleicher Weise argumentiert Alexandra Horowitz in Bezug auf Hunde und Säugetiere allgemein. Sie haben unleugbar Gefühle, aber wir seien nicht in der Lage, deren Andersartigkeit und Komplexität zu beschreiben. Die Hundeforscherin, *Senior Research Fellow* und *Head of the Dog Cognition Lab at Barnard College, Columbia University* postuliert: Wenn sich unser Bewusstsein für kurze Zeit im Körper eines Hundes wiederfände, könnte es nicht die hundlichen Empfindungen und Gefühle in Beziehung setzen zu unseren menschlichen. Das Erfahrene sei zu verschieden. Der entscheidende Kernsatz folgt: „But that there are feelings I've no doubt.“⁴⁷ „As

jeweiligen Kultur, ist somit keine lernbare Reaktion, sondern von Geburt an vorhanden, genetisch disponiert.“

⁴⁵ Neme (2016).

⁴⁶ Ebd.

⁴⁷ Horowitz (2019), S. 201.

a scientist, I don't yet see a way to definitively test another animal's emotional experience."⁴⁸

Was sie aber unter Laborbedingungen zu prüfen vermag, so Horowitz, ist, ob ein bestimmtes tierisches Verhalten, das uns Anlass gibt, es mit einem Attribut versehen (freudig, traurig ...) zu wollen, mehr oder weniger in ähnlichen Kontexten auftritt⁴⁹.

„While the similarities across mammalian brains make it highly likely that all mammals have diverse emotional experiences, we all also have very different lived experiences, based on, for humans, our cultures, where we live, and the people we meet. So, too, for dogs.“⁵⁰

Es ist anzunehmen, dass dies in gleicher Weise auf Grizzlybären zutrifft.

10. Sinn für Ästhetik

Grizzlybären suchen häufig erhöhte Plätze auf, wie Felsen, Berggipfel, und blicken von dort stundenlang entspannt in die Landschaft. Handelte es sich um einen Menschen, hätten wir keine Mühe, diesen oder einen ähnlichen Satz abzuleiten: Er/sie geht auf im Erleben von Schönheit und besitzt daher möglicherweise Sinn für Ästhetik.

„Some believe that bears have the ability to interpret some sense of natural beauty. Scientists have observed bears sitting at vista points for hours staring at a river or mountain views. Coincidence? Possibly.“⁵¹

⁴⁸ Ebd.

⁴⁹ Ebd., S. 208.

⁵⁰ Ebd., S. 201.

⁵¹ Unofficialnetworks.com (2018) Fun Facts Regarding The Intelligence Levels Of Bears.

11. Schlussgedanken: Bären im Anthropozän

Obwohl Grizzlybären hochintelligent und anpassungsfähig sind, erfahren sie oft ein grausames Schicksal aus unserer Hand. Viele Bären werden unnötig gefangen und erschossen. Die Gründe sind Angst und Nachlässigkeit. Sogenannte „Problem-Bären“, die im Hausmüll wühlen oder Gärten oder Höfe durchschnüffeln, sind nur dann ein Ärgernis, wenn Menschen die in *Bear Country* nötigen Vorsichtsmaßnahmen nicht treffen. Beispielsweise: Abfälle sind möglichst geruchsdicht in Garage oder Nebengebäuden aufbewahren; Hühnerställe und Bienenstöcke sind mit Elektrozäunen abzusichern; Hundefutter soll nicht in Hof und Garten aufbewahrt werden.

Da Menschen mehr und mehr Gebiete besiedeln, die traditionell Bären-Territorium sind, kommt es unweigerlich zu Begegnungen und Auseinandersetzungen. So intelligent Bären auch sind, wir Menschen (zu mindestens einige von uns) vermögen in größeren Zusammenhängen zu denken und weitreichende ethische Entscheidungen zu treffen. Koexistenz mit Bären ist möglich und wurde erfolgreich Jahrhunderte lang von indigenen Völkern praktiziert. Die Lernfähigkeit von Bären lässt auch von deren Seite Anpassungen erwarten. Bringen wir ihre Intelligenz mit einem Zitat auf den Punkt: Grizzlybär-Biologe Rob Wielgus, Alberta, behauptet: „The average grizzly bear is always smarter than your average grizzly bear researcher.“⁵²

Literatur

- Ballard, Jack (2018): *Large Mammals of the Rocky Mountains*, Lanham
- Bentham, Jeremy (1789), *An introduction to the principles of mores and legislation*, in: *Animal Rights Philosophy*, [speakingofresearch.com, Zugriff: 14.01.2021]
- Busch, Robert (2000): *The Grizzly Almanac*, Markham
- Craighead, Frank (1982): *Track of the Grizzly*, San Francisco
- Eaton, Brian (2020), *Yukon's Indigenous Peoples and Grizzly Bears*, in: www.grizzlybearprotectionyukon.com, Zugriff: 14.01.2021
- Feddersen-Petersen, Dorit (2008): *Ausdrucksverhalten beim Hund. Mimik und Körpersprache, Kommunikation und Verständigung*, Stuttgart

⁵² Busch (2000), S. 94.

- Grizzly Bear Foundation (2019): *Surviving the winter. What we can learn from hibernating bears*, [Grizzlybearfoundation.com, Zugriff: 14.01.2021]
- Horowitz, Alexandra (2019): *Our dogs, Ourselves: The story of a singular bond*, New York
- Hume, David, *Of the reason of Animals*, in: Animal-Rights-Library.com, Zugriff: 14.01.2021
- Kodiak Bear Center Alaska (2019), www.kodiakbearcenter.com, Zugriff: 14.01.2021
- Morgan, Chris (2008): *Bears of the Last Frontier: Arctic Wanderers, Bear Intelligence*, in: Nature, [www.pbs.org, Zugriff: 14.01.2021]
- Murie, Adolph (1981): *The Grizzlies of Mount McKinley*, US Department of the Interior, Scientific Monograph Series, No 14
- Neme, Laurel (2016): *Remembering the Woman Who Helped Bears in Distress*, in: National Geographic (20.04.2016), [https://www.nationalgeographic.com/animals/article/160420-bears-else-poulsen-animal-emotions-zoos, Zugriff: 14.01.2021]
- Olsen, Lance (1990): Cognitive Complexity of the Grizzly Bear, in: *The Trumpeter, Journal of Ecosophy* 7/4, [http://trumpeter.athabascau.ca/index.php/trumpet/article/view/473, Zugriff: 14.01.2021]
- Regan, Tom/Singer, Peter, Hg. (1989): *Animal Rights and Human Obligations*, Prentice-Hall, New Jersey
- Unofficialnetworks.com (2018): *Fun Facts Regarding The Intelligence Levels Of Bears*, Zugriff: 14.01.2021
- Voigt, Uwe (2021): Künstliche Intelligenz im Anthropozän? Aber natürlich!, in *vorliegendem Band*, S. 109–132

Verzeichnis der Autorinnen und Autoren

Dr. ANNETTE BELKE, Grizzly Bear Protection Yukon. Forschungsschwerpunkt: Praxis und Spiritualität des Tierschutzes. Veröffentlichung: *Lama Anagarika Govinda. Wegbereiter eines „schöpferischen“ Buddhismus im Westen und Begründer des Ordens Arya Matreya Mandala*, Diss. Universität Wien 1995.

Prof. Dr. DIETRICH DÖRNER, Otto-Friedrich-Universität Bamberg. Forschungsschwerpunkt: systemtheoretische Modellierung von Denken als Problemlösen. Wichtige Publikationen: *Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen*, Reinbek bei Hamburg 1989 (¹²2003); *Bauplan für eine Seele*, ebd. 2001; *Die Mechanik des Seelenwagens*, Bern 2002.

Dr. MARION FRIEDRICH, Universität Augsburg. Forschungsschwerpunkt: interdisziplinäre Ansätze in der Neurophysiologie, Psychologie und Gehirnforschung. Wichtige Publikationen: *Die künstliche Evolution der Cyborgs. Erkenntnistheoretische Aspekte der Bioinformatik*, Marburg 2003; *Vom Neuron zum Qubit. Auf den Spuren des Bewusstseins*, Marburg 2008; Wenn Worte nicht mehr reichen: Mit neuroaffektivem Psychodrama von der posttraumatischen Belastung zur posttraumatischen Heilung, in: *Freie Psychotherapie*, H. 2 (2021) [<https://www.vfp.de/magazine/freie-psychotherapie/alle-ausgaben/heft-02-2021/wenn-worte-nicht-mehr-reichen-mit-neuroaffektivem-psychodrama-von-der-posttraumatischen-belastung-zur-posttraumatischen-heilung>, Zugriff: 11.5.2021].

Dr. THOMAS HEICHELE, Universität Augsburg. Forschungsschwerpunkt: Zusammenhang von Technikphilosophie, Wissenschaftstheorie, Naturphilosophie und Künstlicher Intelligenz. Wichtige Veröffentlichungen: *Die erkenntnistheoretische Rolle der Technik bei Leonardo da Vinci und Galileo Galilei im ideengeschichtlichen Kontext*, Münster 2016; Geist-Erfahrung im Wirken Leonardo da Vincis. Der besondere Akt des Zeichnens als zentrales Erkenntnisinstrument, in: Manfred Nege-

le/Jan Levin Propach, Hg., *Geist-Erfahrung. Ein Beitrag zu einem Erfahrungsbegriff für die Geisteswissenschaften*, Würzburg 2019, S. 177–192; als Hg.: *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster 2020.

Prof. Dr. SEAN J. McGRATH, Memorial University of Newfoundland und For a New Earth. Forschungs- und Arbeitsschwerpunkt: Umweltphilosophie vor dem philosophiegeschichtlichen und systematischen Hintergrund des Deutschen Idealismus und der neueren Entwicklungen der Phänomenologie; ökologische Konversion. Wichtige Publikationen: *Heidegger – a (Very) Critical Introduction*, Grand Rapids, MI, 2008; *The Dark Ground of Spirit. Schelling and the Unconscious*, London 2021; *Thinking Nature. An Essay in Negative Ecology*, Edinburg 2019.

Prof. Dr. MICHAEL J. MEITNER, University of British Columbia. Forschungsschwerpunkt: Ästhetische und ökologische Aspekte der Naturwahrnehmung. Wichtige Veröffentlichungen, jeweils als Co-Autor: People's knowledge and risk perceptions of invasive plants in Metro Vancouver, British Columbia, Vancouver, in: *Environmental Management* 66 (2020), S. 985–996; Characteristics of urban green spaces in relation to aesthetic preference and stress recovery, in: *Urban Forestry & Urban Gardening* 41 (2019), S. 6–13; Urban Woodland Understory Characteristic in Relation to Aesthetic and Recreational Preference, in: *Urban Forestry & Urban Gardening* 24 (2017), S. 55–61.

Prof. Dr. UWE MEIXNER, Universität Augsburg. Forschungsschwerpunkt: Metaphysik des Bewusstseins als der Stätte vernünftiger Erkenntnis und freier Entscheidung. Wichtige Veröffentlichungen: *The Two Sides of Being. A Reassessment of Psycho-Physical Dualism*, Paderborn 2004; *Defending Husserl. A Plea in the Case of Wittgenstein & Company Versus Phenomenology*, Berlin–Boston 2014; *Liebe und Negativität*, Münster 2017; *Metaphysik ohne Vorurteile*, Darmstadt 2021.

Dr. JOACHIM RATHMANN, Universität Augsburg. Forschungsschwerpunkt: Ökosystemleistungen, Mensch-Umweltbeziehungen, Therapeutische Landschaften. Wichtige Publikationen: Von der Naturkunde zur Umweltugendethik: Ein möglicher Weg zur Überwindung der Diskrepanz von Umweltwissen und Umwelthandeln?, in: *Comenius-Jahrbuch* 28 (2020) S. 97–120, *Therapeutische Landschaften. Landschaft und Gesundheit in interdisziplinärer Perspektive*. Wiesbaden 2020; mit

Co-Autoren: Using the visitoremployed photography method to analyse deadwood perceptions of forest visitors: a case study from Bavarian Forest National Park, Germany, in: *European Journal of Forest Research* (2020).

Dr. SEBASTIAN ROSENGRÜN, CODE University of Applied Sciences Berlin. Forschungsschwerpunkt: Künstliche Intelligenz und Philosophie digitaler Technik. Wichtige Veröffentlichungen: *Kripkes Metaphysik Möglicher Welten*, Berlin 2019 (als Sebastian Krebs); *Künstliche Intelligenz zur Einführung*, Hamburg 2021.

PD Dr. JENS SOENTGEN, Universität Augsburg. Forschungsschwerpunkt: Phänomenologie der stofflichen Umwelt. Wichtige Veröffentlichungen: *Das Unscheinbare. Phänomenologische Beschreibungen von Stoffen, Dingen und fraktalen Gebilden*, Berlin 1997; *Ökologie der Angst*, Berlin 2018; *Konfliktstoffe. Über Kohlendioxid, Heroin und andere strittige Substanzen*, München 2019.

PD. Dr. STEFANIE VOIGT, Universität Augsburg. Forschungsschwerpunkt: Interdisziplinäre Ästhetiktheorie und ihre Anwendungen. Wichtige Publikationen: *Das Geheimnis des Schönen. Über menschliche Kunst und künstliche Menschen oder: Wie Bewusstsein entsteht. Mit einem Vorwort von Dietrich Dörner*, Münster 2005; *Erhabenheit. Über ein großes Gefühl und seine Opfer*, Würzburg 2011; *Angewandte Ästhetik für Einsteiger. Über Smart Humanities und den neuen (oder alten) Anspruch an die Führung der Industrie von morgen*, Deggendorf 2021.

Prof. Dr. UWE VOIGT, Universität Augsburg. Forschungsschwerpunkt: Philosophie der geistigen Umwelt im Anthropozän. Wichtige Publikationen: Wie ist es, ein Mikrosubjekt zu sein?, in: Manfred Negele/Jan Levin Propach, Hg., *Geist-Erfahrung. Ein Beitrag zu einem Erfahrungsbegriff für die Geisteswissenschaften*, Würzburg 2019, S. 61–78; Wissen um Atmosphären – Bildung für das Anthropozän?, in: *Comenius-Jahrbuch* 28 (2020), S. 13–32; Das Anthropozän als geistige Umweltkrise, in: Thomas Heichele, Hg., *Mensch-Natur-Technik. Philosophie für das Anthropozän*, Münster 2020, S. 85–102.

Personenregister

- Abduljabbar, Rusul 169, 175
Abram, David 248, 251, 267, 273
Achtner, Wolfgang 131
Adorno, Theodor 201
Alberti, Leon Battista 206
Al-Chalabi, Malek 260, 273
Alexy, Robert 204, 212
Al-Farabi 63
Allen, Collin 36, 52
AlphaGo 58
Amelang, Manfred 145, 161
Anaxagoras 59, 72
Anderson, Michael 158, 161
Anscombe, Gertrude Elizabeth Margaret 77
Antonovsky, Aaron 249
Aristoteles 16, 43, 57, 59, 63, 66, 69, 73, 81, 90, 92, 95, 122f., 146, 206, 306
Averroes 63
Avicenna 63
Bachelard, Gaston 279, 292
Bachtin, Michail Michailowitsch 201
Bacon, Francis 93, 95, 262
Ballard, Jack 298, 301, 309
Balzac, Honoré de 178, 180, 184, 197
Bandura, Albert 239, 251
Banse, Gerhard 107
Barck, Karlheinz 212f.
Bartels, Andreas 106
Barthes, Roland 201
Bartneck, Christoph 161
Bartussek, Dieter 145, 161
Bataille, Georges 201
Bateson, Gregory 201, 206
Batthyány, Alexander 107, 265, 272f.
Baumgartner, Christoph 256, 273
Beck, Heinrich 129
Beck, Simon R. 190, 197f.
Beckermann, Ansgar 85, 98, 104
Behler, Ernst 77
Belke, Annette 113, 129
Belton, Valerie 261, 273
Bendel, Oliver 254, 273
Benjamin, Walter 47, 51
Bentham, Jeremy 306, 309
Berlinski, David 81, 104
Bergthaller, Hannes 119, 130
Berry, Wendell 266, 273
Bertalanffy, Ludwig von 91
Betancourt, Michael 47, 51
Beuys, Joseph 201
Beyer, Reinhard 201
Binet, Alfred 147f.
Bird, Geoffrey 250f.
Birnbacher, Dieter 101, 104, 273
Bischoff, Alena 100, 104
Bloom, Paul 264f., 273

- Boileau, Nicolas 207
 Booley 64f.
 Borner, Marc 112, 129
 Bostrom, Nick 55, 58, 75, 77, 83f., 102, 104, 121, 130
 Brandstätter, Ursula 200, 212
 Breithaupt, Fritz 264f., 268, 273
 Breitkopf, Alfred 23
 Brüntrup, Godehard 85, 98, 100, 104f., 131
 Buber, Martin 187, 197
 Buchanan, Bruce G. 81, 105
 Bucher, Anton A. 249, 251
 Buchheim, Thomas 122, 130
 Burke, Edmund 206
 Busch, Robert 296–298, 302f., 309
 Cairns, Dorion 76
 Canli, Turhan 151, 161
 Capen, James 297
 Cardwell, Donald 93, 105
 Carson, Rachel 256
 Cassirer, Ernst 62, 65f., 75, 87–89, 103, 105, 114, 127
 Cattell, Raymond Bernard 150, 161
 Chalmers, David 61, 71, 75, 85, 105
 Charlier, Philippe 248, 251
 Chesterton, Gilbert Keith 75
 Cho, Kyong S. 247, 251
 Chomsky, Noam 201
 Church, Alonzo 22, 26
 Cézanne, Paul 201
 Cicero, Marcus Tullius 146
 Cissna, Kenneth N. 197
 Claessens, Dieter 115, 130
 Colli, Giorgio 274
 Cook, Jill 67, 75
 Cook, Steven J. 165, 175
 Coolidge, Frederick L. 67, 77
 Craighead, Frank 298, 304, 309
 Craighead, John 298
 Crevier, Daniel 175
 Crowe, Frederick E. 76
 Crutzen, Paul J. 119, 130
 Cusanus, Nicolaus 93, 95
 Cyffka, Bernd 255, 274
 Darwin, Charles 306
 Data 199, 203–205, 211
 Davis, Martin 81, 105, 116, 130
 Dawkins, Richard 72
 Deep Blue 20, 58
 Dennett, Daniel 35f., 51
 Descartes, René 7, 124f., 305
 Devall, Bill 266, 273
 Diekmann, Andreas 236
 Dionysios Areopagita 211
 Dogs, Christian Peter 161
 Doran, Robert M. 76
 Dörner, Dietrich 113, 118, 130, 204, 212, 233
 Drachman, David A. 165, 175
 Draisma, Douwe 204, 212
 Dresler, Martin 84, 102, 105
 Dreitzel, Hans P. 256, 273
 Dreyfus, Hubert 41, 51, 58–61, 64, 71, 75
 Duerr, Hans Peter 201
 Dufresne, Todd 120, 130
 Dujon, Antoine M. 168, 176
 Duns Scotus, Johannes 63

- Eaton, Brian 296, 309
 Edmonds, Dean 164
 Einstein, Albert 16f., 210
 ELIZA 164
 Engel, Gisela 87, 105
 English, Jennifer 251
 Epstein, Robert 176
 Ernst, Max 209
 Eusterschulte, Anne 106
 Eysenck, Hans Jürgen 137, 142, 151, 156, 161
 Fechner, Gustav 63, 75
 Feddersen-Petersen, Dorit 306, 309
 Fencil, Kevin 137, 177
 Festinger, Leon 201, 236, 256, 273
 Fichte, Johann Gottlieb 61, 76, 112–114, 130
 Ficino, Marsilio 206
 Fingerhut, Joerg 98, 105
 Fink, Bruce 76
 Fjelland, Ragnar 57f., 60, 76
 Fliessbach, Klaus 197
 Floridi, Luciano 94, 105
 Flusser, Vilém 201
 Flynn, James R. 155, 161
 Forrester, Frank 77
 Franke, Ursula 202, 212
 Frankish, Keith 35, 51
 Frankl, Viktor E. 241, 244, 249, 252
 Franz, Michael 200, 212
 Freeland, Cynthia 63, 76
 Frege, Gottlob 16–22, 24, 81
 Friedrich, Marion 111, 129f.
 Fuchs, Thomas 262f., 273
 Fulton, Scott 165, 176
 Funke, Joachim 146, 161
 Funke, Pia-Maria 209, 212
 Gabriel, Markus 60f., 76
 Gadamer, Hans-Georg 201, 280
 Gaia 119
 Galilei, Galileo 93, 96, 129
 Gall, Franz Joseph 146
 Gallee, Martin A. 81, 87f., 95f., 101, 105
 Galton, Francis 147
 Gardener, Howard 143, 161
 Gebauer, Sebastian 291f.
 Gehlen, Arnold 89, 96, 105, 256, 273
 Geier, Fabian 33
 Gesang, Bernward 45, 51
 Gittler, Georg 162
 Glasser, Ira 173, 176
 Göbel, Richard 94, 108
 Goleman, Daniel 152, 161, 266, 273
 Gonzales, Robert 190, 198
 Good, Irving J. 54f., 84, 105
 Gorke, Michael 268f., 273
 Görz, Günther 97, 105
 Gottl-Ottlilienfeld, Friedrich von 86f., 89, 105
 Greve, Jens 100, 105
 Grizzly Adams: *siehe* Capen, James
 Grunwald, Armin 82, 105, 107, 281, 292
 Grupe, Gisela 92, 106
 Güclü, Nurten Güler 9
 Guilford, Joy Paul 150
 Habermas, Jürgen 122, 130
 Haenlein, Michael 140, 161
 HAL 55
 Halla, Michael 257, 273

- Hammermeister, Kai 201, 212
 Hansen, Margaret 246, 252
 Harari, Yuval Noah 229, 233
 Harbisson, Neil 45, 51
 Harris, Marvin 201
 Harris, Roy 77
 Hartmann, Nicolai 265f., 273
 Hawking, Stephen 139, 162
 Hayden-Roy, Priscilla 77
 Hebb, Donald O. 97
 Hegel, Georg Wilhelm Friedrich 61, 63, 76
 Heichele, Thomas 33, 80–82, 86–90, 92–96,
 98, 102–104, 106, 108, 114, 119, 129f., 131f.
 Heidegger, Martin 41, 58f., 64, 73, 76, 116, 130
 Heilmann, Till 46, 51
 Hengstenberg, Hans-Eduard 270f., 273
 Henke, Winfried 92, 106,
 Henrich, Dieter 113, 130
 Hephaistos 254
 Herculano-Houzel, Suzana 165, 176
 Herpertz, Sabine 197
 Herrigel, Eugen 76
 Herrmann, Friedrich-Wilhelm von 76
 Hirdina, Karin 212
 Hollingdale, Reginald John 77
 Homer 254
 Hope, Cat 178, 180, 183, 186
 Hopfield, John 96
 Horn, Eva 119, 130
 Horowitz, Alexandra 307f., 310
 Hösle, Vittorio 112, 130
 Hoyningen-Huene, Paul 100, 106
 Hubig, Christoph 82, 86f., 89, 106
 Hufendiek, Rebekka 98, 105
 Hume, David 306, 310
 Husserl, Edmund 63, 69, 73, 76
 Illies, Christian 80, 103, 106
 Immich, Gisela 245f., 252
 Ingensiep, Hans W. 106
 Irle, Martin 143, 161
 Irrgang, Bernhard 40, 115, 131
 James, William 201
 Janich, Peter 86, 106
 Jaskolla, Ludwig 100, 105, 131
 Jaynes, Julian 201, 204, 212
 Julliard, Yannick 82, 105
 Julmi, Christian 132
 Jung, Carl Gustav 62
 Jungert, Michael 281, 292
 Kaden, Christian 206, 212
 Kant, Immanuel 209, 212, 250, 252
 Kaplan, Andreas M. 140, 161
 Kaplan, Jerry 83, 85, 106
 Kaplan, Stephen 246, 252
 Kaplan, Rachel 246, 252
 Kapp, Ernst 82, 95, 103, 106
 Karafyllis, Nicole C. 87, 90, 106
 Kekulé, August 278
 Kelsey, Henry 296
 Kirchhoff, Thomas 90, 106
 Kirsch, Peter 197
 Kishon, Ephraim 202, 212
 Kittler, Friedrich 46, 51
 Klages, Ludwig 255, 256, 273
 Klingan, Kartin 119, 131
 Koehler, Nicole 182

- Kolany-Raiser, Barbara 94, 106
 Köller, Wilhelm 282f., 292
 Konrath, Sara H. 274
 Korpela, Kalevi M. 249, 252
 Kornwachs, Klaus 81, 106
 Koyré, Alexandre 125, 131
 Kraus, Marita 132, 197
 Krebs, Angelika 268, 274
 Krebs, Sebastian 38f., 43, 51
 Kripke, Saul Aaron 38f., 43, 51, 126f., 131
 Krohn, Wolfgang 86, 106, 200, 212, 281, 292
 Kuckartz, Udo 236, 238, 252
 Kuhlmann, Meinhard 91, 106
 Kuhn, Helmut 262, 274
 Kurzweil, Ray 55, 178f., 183, 197
 Kutschera, Franz von 23
 Lacan, Jacques 66, 76
 Lamberti, Georg 147, 161
 Langer, Susanne K. 63, 65f., 76
 Lask, Emil 73, 76
 Latour, Bruno 117, 131
 Lauricella, Tom 170, 176
 Le Goff, Jacques 203
 Leibniz, Gottfried Wilhelm 22, 81
 Leinfelder, Reinhold 94, 106
 Lenzen, Manuela 83, 85, 94f., 98, 101, 103, 106
 Leonardo da Vinci 93, 95, 96, 103, 201
 Lessing, Gotthold Ephraim 206, 208
 Levy, David 43, 51
 Lichtenberg, Georg Christoph 278f., 282, 284f., 291f.
 Liggieri, Kevin 51
 Lindemann, Wilhelm 132
 Lipps, Theodor 188, 197, 265, 274
 List, Elisabeth 122, 126, 131
 Loh, Janina 40, 43, 45, 51, 84, 106
 Lohse, Simon 108
 Lonergan, Bernard 57, 63, 65, 69, 72f., 76
 Lovelock, James 119, 121, 131
 Lübke, Hermann 257, 259, 262, 274
 Luhmann, Niklas 281, 292
 Łukasiewicz, Jan 17–22, 24
 Lurija, Alexander Romanowitsch 201
 Lyotard, Jean-François 201
 Lyre, Holger 98, 106
 MacGregor, Neil 67, 76
 Macquarrie, John 76
 Maier, Steven F. 242, 252
 Main, Mary 156, 161
 Mainzer, Klaus 35f., 40, 51, 81, 84f., 91, 94, 96–98, 100–102, 107, 116, 121, 131
 Malewitsch, Kasimir Sewerinowitsch 201
 Mangal, S K 161
 Markl, Hubert 289, 292
 Marttunen, Mika 274
 Matthews, Freya 131
 Maturana, Humberto 210–212
 Mayer, John D. 143, 150, 153, 161
 McCarthy, John 85
 McGinn, Colin 99, 107
 McGovern, Amy 169, 176
 McGrath, Sean J. 7f., 65, 74, 76, 118, 122, 125, 131
 Meitner, Michael J. 121, 131
 Meixner, Uwe 9, 99, 107, 112, 118, 127, 129, 131
 Menninghaus, Winfried 206, 210, 212

- Merton, Thomas 271, 274
- Metzinger, Thomas 36, 48, 49, 52, 85, 98, 102, 104, 107
- Mikat, Paul 292
- Milner, Peter 194, 197
- Minsky, Marvin 57, 85, 164, 176
- Misselhorn, Catrin 40, 52
- Mittelstraß, Jürgen 87, 90, 105–107, 281f., 292
- Mohler, George 96, 107
- Moor, James H. 136, 138, 142, 157f., 161
- Moran, Dermont 76
- Moravec, Hans 141, 162
- Morgan, Chris 302f., 310
- Moscovici, Serge 122
- Müller, Michael 130
- Müller, Oliver 51
- Müller, Vincent 58, 77
- Müller-Röber, Bernd 92, 107
- Murie, Adolph 303, 310
- Musk, Elon 139, 162
- Mustajoki, Iyri 274
- Nachtigall, Werner 96, 107
- Nordmann, Alfred 81f., 86, 107
- Nagel, Thomas 60, 61, 70–74, 77
- Nassehi, Armin 47, 52
- Naterer, Greg 58
- Needham, Joseph 279
- Negele, Manfred 104
- Neme, Laurel 307, 310
- Neumann, Gerhard 282, 292
- Newton, Isaac 17
- Nietzsche, Friedrich 72, 77, 204, 265f., 274
- Nilsson, Nils 35, 52
- Norvig, Peter 35, 52, 81, 83, 97, 103, 107
- Oehler-Klein, Sigrid 146, 162
- Olds, James 194, 197
- Olsen, Lance 297f., 301, 304f., 310
- Ong, Walter J. 201
- Ott, Maximilian 94, 107
- Otto, Rudolf 247, 252
- Panofsky, Erwin 201
- Pape, Helmut 113, 131
- Peirce, Charles Sanders 63, 65, 77, 201, 222, 233
- Peitsch, Helmut 204, 213
- Peschl, Markus F. 107
- Peterson, Christopher 252
- Picasso, Pablo 201
- Pietsch, Wolfgang 94, 107
- Pietschnig, Jakob 155f., 162
- Platon 63, 205f.
- Plotin 63, 211
- Plumwood, Val 126, 131
- Polanyi, Michael 58, 77
- Popitz, Heinrich 115, 131f.
- Porges, Stephen W. 244, 252
- Poulsen, Else 307, 310
- Propach, Jan Levin 104
- Pühler, Alfred 92, 107
- Putnam, Hilary 38, 52, 102, 107
- Quine, Willard van Orman 43, 52
- Raabe, Wilhelm 255
- Ramsey, William 35, 51
- Rapoport, Anatol 91
- Rathmann, Joachim 121, 129f., 132, 246, 248, 252, 257, 274

- Recki, Birigt 105
 Regan, Tom 305, 310
 Reicher, Maria Elisabeth 200–202, 213
 Renn, Jürgen 119, 132
 Reschke, Renate 212
 Reydon, Thomas 108
 Ricœur, Paul 62, 66, 77
 Rigby, Kate 131
 Rizzolatti, Giacomo 153, 162
 Robinson, Edward 76
 Rochester, Nathaniel 85
 Rogers, Carl 188, 193f., 197
 Rolston, Holmes III 269, 274
 Ropohl, Günter 82, 86, 91, 107
 Rose, Deborah Bird 131
 Rosenberg, Jay F. 80, 107
 Rosenberg, Marshall B. 187, 197
 Rosengrün, Sebastian 34f., 37f., 46f., 49–52, 80f., 83–85, 92, 97f., 101f., 107, 117f., 132
 Rosol, Christoph 119, 131
 Rost, Detlef H. 162
 Rothe, Hartmut 92, 106
 Rousseau, Jean-Jacques 278, 284
 Russell, Bertrand 38
 Russell, Stuart 35, 52, 81, 83, 97, 103, 107
 Russer, Achim 131
 Ryle, Gilbert 70
 Sachsse, Hans 87f., 107, 114
 Salovey, Peter 143, 150, 152f., 161f.
 Sánchez Bodas, Ricardo Andres 187, 197
 Saß, Henning 197
 Saussure, Ferdinand de 66, 77
 Savage, Denis 77
 Schelling, Friedrich Wilhelm Joseph 61, 77
 Schelsky, Helmut 288f., 292
 Scherer, Bernd 119, 132
 Schiemann, Gregor 122–124, 132
 Schiller, Friedrich 207
 Schmaus, Thomas 121, 132
 Schmid, Ute 97, 105
 Schmidt, Matthias 121, 132
 Schnabel, Annette 100, 105
 Schneider, Kurt 197
 Schofield, Gail 168, 176
 Schöne-Seifert, Bettina 45, 52
 Schrems, Isabel 257, 274
 Schrott, Raul 254, 274
 Schuh, Angela 245f., 252
 Schurz, Gerhard 91, 102, 108
 Schwab, Klaus 93, 108
 Schwibs, Bernd 131
 Searle, John 60f., 77, 83f., 108
 Sedlmayr, Hans 201
 Seligman, Martin E. P. 239, 241f., 252
 Seus, Doug 297f.
 Shamay-Tsoory, Simone G. 195, 197
 Shannon, Claude 85
 Shannon, Lorraine 131
 Siep, Ludwig 178, 180, 184, 197
 Simonins, Georg 292
 Singer, Peter 305, 310
 Sinigaglia, Corrado 162
 Skinner, Burrhus F. 191, 197
 Sloterdijk, Peter 117, 132, 201
 Smith, Brian Cantwell 60, 77

- Soentgen, Jens 62, 77, 90, 108, 117, 119f., 122, 126f., 132, 180, 197, 252, 274, 292
- Sokolowski, Robert 74, 77
- Sokrates 59f., 72
- Spearman, Charles 146, 149
- Spitzer, Manfred 155, 162, 195, 198, 250, 252
- Spranger, Eduard 263, 274
- Städte, Klaus 201, 213
- Stammel, Barbara 255, 274
- Steffen, Will 258, 274
- Stendhal 206
- Stephan, Achim 103, 106, 108
- Stewart, Theo 261, 273
- Stöckler, Manfred 106
- Stern, William 136f., 140, 146, 148f., 159, 162
- Sturm, Thomas 103, 108
- Szymborska, Wislawa 264, 274
- Talaviya, Tanha 168, 176
- Talbot, Davinia 45, 52
- Taylor, Charles 65, 77
- Tegmark, Max 50, 54, 77, 83, 85, 98, 102, 108
- Tetens, Holm 99, 108
- Thomas von Aquin 62f.
- Thompson, Evan 60, 77
- Thorndike, Edward Lee 143, 150, 162
- Thurstone, Louis Leon 150
- Tillich, Paul 62, 66, 77
- Timpany, Phil 296, 299
- Tomasello, Michael 92, 108
- Tredennick, Hugh 60
- Turing, Alan 61, 69, 81, 108, 164, 176
- Twenge, Jean M. 264, 274
- Ukowitz, Martina 281, 292
- Upham, Paul 257, 274
- Varela, Francisco J. 210–212
- Vaterrodt, Bianca 146, 161
- Vetter, August 266, 274
- Voigt, Stefanie 113, 123, 127, 132, 200–202, 204, 206f., 210, 213
- Voigt, Uwe 33, 36, 51f., 56, 62, 77, 94, 100, 108, 120, 123f., 128, 132, 281, 292, 295, 310
- Waal, Frans de 264, 275
- Wachsmuth, Ipke 97, 105
- Wallace, William 76
- Wallach, Wendel 36, 52
- Walter, Sven 103, 106, 108
- Wandruszka, Mario 284, 286, 292
- Wechsler, David 143, 148, 150
- Weitze, Marc-Dennis 92, 107
- Weizenbaum, Joseph 35, 52, 164, 176
- Weizsäcker, Carl Friedrich von 287, 292
- Welding, Steen Olaf 102, 108
- Welsch, Wolfgang 92, 108, 202, 213
- Wernecke, Jörg 94, 107
- White, Matthew P. 245, 252
- Whitehead, Alfred North 63
- Whorff, Benjamin Lee 279
- Wielgus, Rob 309
- Williams, Rowan 72
- Wirtz, Markus A. 146, 162
- Wittgenstein, Ludwig 61, 70, 77, 135, 137, 142, 145, 162, 201, 207, 268, 275
- Wolf, Barbara 132
- Wolf, Stefan 52
- Wolf, Ursula 138
- Wolff, Dietmar 94, 108
- Wong, Zoie S. Y. 168, 176

Wongupparaj, Peera 154, 155, 162

Wright, William 304

Wynn, Thomas 67, 77

Zapf, Hubert 121, 132

Ziegler, Dieter 93, 108

Zimmerli, Walther Ch. 52

Zoglauer, Thomas 81f., 86, 96, 108

Zorn, Daniel-Pascal 122, 127, 132

